Generalized Symplectic Algorithms for Differential Equations

微分方程的广义辛算法

苏红玲　秦孟兆◎著



北京大学出版社
PEKING UNIVERSITY PRESS

微分方程的广义辛算法

苏红玲　理学博士、中国人民大学信息学院教师、硕士生导师。主要研究领域为动力系统保结构算法。分别于 1997 年和 2000 年在兰州大学取得计算数学专业理学学士和应用数学专业理学硕士学位，并于 2003 年在中国科学院数学与系统科学研究院取得计算数学专业理学博士学位。作为访问学者，曾访问德国 Karlsruhe 大学和美国 Los Alamos 国家实验室。曾主持和参与国家自然科学基金青年基金项目和面上项目。

秦孟兆　1961 年毕业于莫斯科大学数学力学系。中国科学院数学与系统科学研究院研究员、博士生导师．主要研究领域为动力系统保结构算法。作为冯康科技组的成员，所参与的项目"哈密尔顿辛几何算法"获 1997 年国家自然科学一等奖。曾主持及参与多项国家自然科学基金项目。出版的著作有《哈密尔顿系统的辛几何算法》《偏微分方程中的保结构算法》（获第三届中国出版政府奖图书奖）和 *Symplectic Geometric Algorithms for Hamiltonian Systems*（因保持高收费下载量获 Springer 出版社年报通报）。



"北京大学出版社"
微信公众号

上架建议：理工
ISBN 978-7-301-25996-2



定价：32.00元

Generalized Symplectic Algorithms for Differential Equations

微分方程的广义辛算法

苏红玲　秦孟兆◎著



北京大学出版社
PEKING UNIVERSITY PRESS

图书在版编目 (CIP) 数据

微分方程的广义辛算法 / 苏红玲，秦孟兆著 . -- 北京 : 北京大学出版社，2015.7
ISBN 978-7-301-25996-2

Ⅰ. ①微… Ⅱ. ①苏… ②秦… Ⅲ. ①微分方程 - 高等学校 - 教材 Ⅳ. ① O175

中国版本图书馆 CIP 数据核字 (2015) 第 143636 号

书　　名	微分方程的广义辛算法
著作责任者	苏红玲　秦孟兆　著
责任编辑	曾琬婷
标准书号	ISBN 978-7-301-25996-2
出版发行	北京大学出版社
地　　址	北京市海淀区成府路 205 号　100871
网　　址	http://www.pup.cn　新浪微博 : @ 北京大学出版社
电子信箱	zpup@pup.cn
电　　话	邮购部 62752015　发行部 62750672　编辑部 62767347
印 刷 者	北京大学印刷厂
经 销 者	新华书店
	880 毫米 × 1230 毫米　A5　7.25 印张　214 千字
	2015 年 7 月第 1 版　2015 年 7 月第 1 次印刷
定　　价	32.00 元

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有，侵权必究

举报电话：010-62752024　电子信箱：fd@pup.pku.edu.cn

图书如有印装质量问题，请与出版部联系，电话：010-62756370

内 容 简 介

本书从基本概念出发, 结合实际应用, 系统地介绍了 Birkhoff 系统的变分方法、几何结构及其保结构算法. 为了使读者能够循序渐进地理解和掌握 Birkhoff 系统的几何结构和保结构算法的构造, 本书首先介绍内容所涉及的力学、代数、几何和 Hamilton 辛算法的基础知识; 然后给出了包含 Hamilton 系统和 Birkhoff 系统的最一般 Euler-Lagrange 方程 (组) 的判别准则, 并讨论了 Birkhoff 系统对应变分问题的几何变分原理; 最后系统地讲解了有限维 Birkhoff 系统和有限维 Birkhoff 辛结构、Birkhoff 多辛系统和 Birkhoff 多辛结构以及无限维 Birkhoff 系统和无限维 Birkhoff 辛结构. 基于这三种系统和三种结构, 定义了有限维 Birkhoff 辛算法、Birkhoff 多辛算法以及无限维 Birkhoff 辛算法, 并详细讲解了构造这三种算法的方法: 生成函数方法、离散变分方法和生成泛函方法.

我们尽可能系统地介绍了 Birkhoff 系统的保结构算法, 希望这些内容对研究动力系统保结构算法的研究生和科研人员具有一些参考价值.

前　言

辛几何算法自中国数学家冯康和美国数学家 Ruth 提出以来, 发展迅速, 获得了广泛应用. 辛几何算法为有限维 Hamilton 系统提供了一种有效的数值方法. 算法的构造方法包括生成函数方法、辛 Runge-Kutta 方法和组合方法等, 定性分析则包括数值 KAM 理论和后向误差分析等. 在实践中, 辛几何算法在等离子体、电磁场、量子力学、天体力学、分子动力学以及流体动力学等领域得以广泛应用, 也取得了很好的效果. 1988 年, 本书作者之一秦孟兆首次将辛几何算法应用于无限维 Hamilton 系统. 其基本思想是: 首先对空间进行适当的有限差分离散, 使得所得的半离散方程是时间方向上的有限维 Hamilton 系统; 然后利用辛几何算法对有限维 Hamilton 系统进行数值计算. 这种算法保持了无限维 Hamilton 系统的整体辛结构. Bridges 于 1997 年首次提出了无限维 Hamilton 系统的多辛 Hamilton 形式, 此形式在时间方向和每一个空间方向都存在一个辛结构, 并在局部意义上满足一个多辛守恒律. 针对多辛 Hamilton 形式的数值算法主要有多辛 Runge-Kutta 算法等. 继 Bridges 提出多辛形式的概念之后, Marsden 等人利用 Veselov 离散力学的思想发展了一种变分算法, 提出了多辛变分积分子的概念. 和多辛 Runge-Kutta 算法一样, 多辛变分积分子也保持多辛守恒律.

与一般 Hamilton 系统算法方面的发展相比, 对受外力作用和约束的 Hamilton 系统的研究成果比较少. Marsden 等人研究了一类受外力作用的 Lagrange 系统和 Hamilton 系统. 当外力满足一定条件时, 这类系统会保持 Hamilton 辛结构, 此时 Hamilton 辛算法对这类系统是保辛结构的. 但是, 这样相当于把外力作用和辛结构分开处理, 过于简化了外力作用, 而实际应用中, 动力系统的几何结构和外力作用之间是相关的. 我们通过 Birkhoff 系统很自然地结合了两个问题: 受外力作用

的系统和具有一般辛结构的系统.

Birkhoff 系统是 R. M. Santilli[San83a,San83b] 于 1978 年首次提出的. Birkhoff 系统是 Hamilton 系统的推广, 是具有一般辛结构的受外力作用的对称系统. 中国学者梅凤翔等人研究了 Birkhoff 系统的对称性和守恒律[Mei96], 武际可等人研究了没有耗散项的自治 Birkhoff 系统辛算法[ZHW99,ZHW02], 做出了一系列重要工作. 笔者于 2002 年前后提出了 Birkhoff 保结构算法 [SQ04], 并将 Birkhoff 系统及其保结构算法推广到无限维情形 [SQ10,SL14]. 近年来, 出现了很多研究 Birkhoff 算法的文献[ZHW02,LLG11,KWM12,XPZL12,KWM12,KWM113,LSG13,GZH13]. 本书的目的是使读者系统掌握有限维和无限维 Birkhoff 系统的几何结构, 并掌握其保结构算法的构造方法.

全书共八章. 第一章从力学、几何和代数三个方面介绍一些基础知识. 第二章简单介绍了 Hamilton 系统的辛算法. 第三章讨论了不同形式的 Euler-Lagrange 方程的对称性判别条件, 分析了一阶 Euler-Lagrange 方程的对称性, 从而说明系统局部辛结构和所受外力之间的关系. 在第四章中，我们分析了 Pfaff 作用泛函表示的受外力作用的奇异 Lagrange 力学系统的性质, 利用不同的变分方法讨论了 Pfaff 作用泛函问题以及有限维和无限维 Birkhoff 系统的对应关系及系统的几何结构, 并利用离散变分构造了 Birkhoff 系统的保结构算法. 第五章从 Birkhoff 常微分方程出发, 讨论了这种方程的对称性以及它的 Birkhoff 辛结构和 Birkhoff 结构, 给出了保 Birkhoff 辛结构数值格式生成函数构造方法. 第六章介绍了受外力作用的多辛 Birkhoff 形式. 它是 Bridges 的多辛 Hamilton 形式向两个方面的推广: 一是各个方向的 Hamilton 辛结构被推广到一般辛结构; 二是系统受到外力作用. 第七章介绍了无限维 Birkhoff 系统及其保结构算法, 详细讨论了构造无限维 Birkhoff 辛格式的生成泛函方法. 第八章作为应用, 讨论了不同形式的电磁场方程的对称性和多辛形式, 并讨论了其保结构算法. 书中给出了一些数值算例, 并在第五章的附录中给出了利用生成函数构造 Birkhoff 辛格式的详细推导过程.

本书得到了国家自然科学基金 (No. 10701081, 11071251) 的资助, 北京大学出版社对本书的出版给予了极大的支持, 在此一并表示感谢.

本书的引用文献仅代表作者的研究兴趣, 如果漏掉重要的文献, 敬请谅解. 由于作者水平有限, 本书错漏与不足之处在所难免, 请各位读者批评指正.

作 者

2015 年 3 月

目　录

第一章　一些力学、几何、代数的预备知识

在"问题原型的基本特征在离散后应该尽可能地得到保持"的原则指导下, 冯康首先提出"保结构算法"的概念. 他和他的研究小组在 Hamilton 系统辛算法的格式构造和理论分析方面取得了一系列成果, 使得"算法几何化"的思想引起国内外学者的极大兴趣. 所谓的算法几何化, 简单地说就是在构造一个动力系统的数值算法时, 要尽可能地保持系统原有的几何性质. 几何化的算法称为几何积分. 几何积分是微分方程的数值积分, 同时它能精确地保持 (在舍入误差范围内) 方程的一个或多个几何特性, 比如保持能量、动量、角动量、相空间体积或面积、对称性、时间反演对称性、辛结构和耗散性等等. 这些几何特性在物理中是至关重要的, 因此, 要学习动力系统的保结构算法, 首先应该掌握动力系统的力学描述, 几何背景和涉及的代数知识. 本章我们首先介绍一些后面要用到的基础知识, 分别从力学、几何和代数三个角度来讲[San83,Arn89,FQ09].

§1.1　经典力学的预备知识

17 世纪, Newton 对于自然科学的伟大贡献之一就是提出了力学基本定律 $F = ma$. 假设一个有 n 个自由度的运动系统, 它的位置向量为 $\boldsymbol{q} = (q_1, \cdots, q_n)^{\mathrm{T}}$, 速度向量为 $\dot{\boldsymbol{q}} = (\dot{q}_1, \cdots, \dot{q}_n)^{\mathrm{T}}$, 加速度向量为 $\ddot{\boldsymbol{q}} = (\ddot{q}_1, \cdots, \ddot{q}_n)^{\mathrm{T}}$. 设这个系统是一个守恒的力学体系, 它的力场 $\boldsymbol{F}$ 表示为某个势能函数 $V = V(\boldsymbol{q})$ 的负梯度:

$$\boldsymbol{F} = -\frac{\partial}{\partial \boldsymbol{q}} V(\boldsymbol{q}),$$

于是

$$\boldsymbol{m}\ddot{\boldsymbol{q}} = -\frac{\partial}{\partial \boldsymbol{q}} V(\boldsymbol{q}). \tag{1.1.1}$$

这是守恒运动方程的标准形式, 它是 n 维空间 $\mathbf{R}^n$ 中的二阶微分方程组, 通称为经典力学的标准形式, 其中质量 $\boldsymbol{m}(\boldsymbol{q})$ 为对称正定矩阵.

到了 18 世纪, Euler 及 Lagrange 引进了作用量:

$$L = \text{动能} - \text{势能} \triangleq T - V.$$

这就是 Lagrange 函数:

$$L(\boldsymbol{q},\dot{\boldsymbol{q}}) = T(\dot{\boldsymbol{q}}) - V(\boldsymbol{q}) = \frac{1}{2}\langle\dot{\boldsymbol{q}},\boldsymbol{m}\dot{\boldsymbol{q}}\rangle - V(\boldsymbol{q}).$$

他们证明了在给定初始位置 $\boldsymbol{q}(t_0) = \boldsymbol{q}^0$ 和终止位置 $\boldsymbol{q}(t_1) = \boldsymbol{q}^1$ 的一切假想轨道中, 能使泛函 $\displaystyle\int_{t_0}^{t_1} L(\boldsymbol{q},\dot{\boldsymbol{q}})\mathrm{d}t$ 达到极值的轨道就是运动的真实轨道, 真实轨道的 $\boldsymbol{q}$ 满足变分方程

$$\delta\int_{t_0}^{t_1} L(\boldsymbol{q},\dot{\boldsymbol{q}})\mathrm{d}t = 0,$$

其中 δ 是变分微分算子. 事实上, 这个变分原理的极值方程 (Euler-Lagrange 方程)

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \boldsymbol{q}} - \frac{\partial L}{\partial \boldsymbol{q}} = 0 \tag{1.1.2}$$

就等价于 Newton 方程 (1.1.1). 方程 (1.1.2) 称为**经典力学的变分形式**, 即 Lagrange 形式, 是由 n 个二阶常微分方程组成的 Lagrange 方程组.

到了 19 世纪, Hamilton 又提出了另一种方程形式. 引进广义动量 $\boldsymbol{p} = (p_1,\cdots,p_n)^{\mathrm{T}}, p_i = \dfrac{\partial L}{\partial \dot{q}_i}$, 同时对 Lagrange 函数 $L(\boldsymbol{q},\dot{\boldsymbol{q}})$ 作 Legendre 变换, 得到 Hamilton 函数

$$\begin{aligned} H(\boldsymbol{p},\boldsymbol{q}) &= \langle\boldsymbol{p},\dot{\boldsymbol{q}}\rangle - L(\boldsymbol{q},\dot{\boldsymbol{q}}) = \langle\dot{\boldsymbol{q}},\boldsymbol{m}\dot{\boldsymbol{q}}\rangle - L = \langle\boldsymbol{p},\boldsymbol{m}^{-1}\boldsymbol{p}\rangle - L \\ &= \langle\boldsymbol{p},\boldsymbol{m}^{-1}\boldsymbol{p}\rangle - (T - V(\boldsymbol{q})) = \frac{1}{2}\langle\boldsymbol{p},\boldsymbol{m}^{-1}\boldsymbol{p}\rangle + V(\boldsymbol{q}). \end{aligned}$$

由此知道 $\mathrm{d}H = \mathrm{d}(\boldsymbol{p}^{\mathrm{T}}\dot{\boldsymbol{q}}) - \mathrm{d}L = \dot{\boldsymbol{q}}\mathrm{d}\boldsymbol{p} - \dfrac{\partial L}{\partial \boldsymbol{q}}\mathrm{d}\boldsymbol{q}$. 另一方面, $\mathrm{d}H = \dfrac{\partial H}{\partial \boldsymbol{q}}\mathrm{d}\boldsymbol{q} + \dfrac{\partial H}{\partial \boldsymbol{p}}\mathrm{d}\boldsymbol{p}$. 比较这两个关系式得到

$$\dot{\boldsymbol{p}} = -\frac{\partial H}{\partial \boldsymbol{q}}, \quad \dot{\boldsymbol{q}} = \frac{\partial H}{\partial \boldsymbol{p}}. \tag{1.1.3}$$

此方程组称为 **Hamilton 正则方程组**, 它是一阶微分方程组, 具有极其简洁且对称的形式. 其中, n 个自由度 $\boldsymbol{q} = (q_1, \cdots, q_n)$ 构成的空间叫做**位形空间** (configuration space), 记为 M; $(\boldsymbol{p}, \boldsymbol{q})$ 称为**相变量**, 其构成的 $2n$ 维相空间 (phase space) 是一个辛空间 (symplectic space).

Lagrange 方程和 Hamilton 方程虽然等价地刻画了同一物理规律, 但因为形式上的差异, 所以在实际数值计算时是不等效的. 辛算法是以 Hamilton 方程为基本形式的[FQ91a], 因为: 首先 Hamilton 方程具有非常简洁且对称的形式, 运动的规律性在 Hamilton 形式下体现得最明显; 其次, Hamilton 形式相比 Newton 形式更为广泛存在和普适, 它覆盖了经典的、相对论的、量子的、有限或无限自由度的、耗散效应可忽略的物理过程. 所以针对 Hamilton 方程的辛算法研究有广阔的发展和应用.

传统算法中除了个别算法, 多数为非辛算法, 它们大多数是面向渐近稳定系统设计的, 都含有耗散机制以保证计算稳定性. 而 Hamilton 系统不具有渐近稳定性, 所以将传统的非辛算法用于解决 Hamilton 系统时, 这些算法不可避免地带进人为耗散性、虚假吸引子以及其他种种非 Hamilton 系统本身所有的寄生效应, 在长时间计算中将最终导致计算结果严重歪曲和失真. 因此, 传统非辛算法用于短期的、瞬态的模拟效果尚可, 但用于长期跟踪和整体的结构性研究则不行, 会引向错误的结论.

辛算法不含人为耗散性, 先天性地免于一切非辛污染, 是 “干净” 的算法. Hamilton 系统拥有两类守恒律: 一类是相空间内偶数面积的不变性, 即 Liouville-Poincaré 守恒律, 它们在辛算法下被自动保持; 另一类是包括能量在内的运动不变量, 如动量、角动量守恒律等. 另外, 一切辛算法都拥有自己的形式不变量, 它们对于原有不变量的逼近阶与算法本身的逼近阶相当. 相关文献也初步证明了在辛算法下近似可积系统的不变环大部分得到保持, 这是著名的 KAM (Kolmogorov-Arnold-Moser) 定理的一种新模式[Kol54],[Arn63],[Mos62],[Shang99]. 这同时表明, Hamilton 算法的结构体系与守恒律完整并行, 高度靠近于 Hamilton 系统原形, 而且拥有理论上无限长期的跟踪能力. 这就从根本上保证并说明了 Hamilton 算法的优越性. 因此, 一条正确计算 Newton 运

动方程的途径就是先把方程 Hamilton 化, 然后运用 Hamilton 算法.

§1.2 几何的预备知识

1.2.1 微分流形

流形是近代数学中最重要的概念之一. 许多物理问题中会自然地出现拓扑空间. 例如, 图 1.2.1(a) 所示的单摆运动, 它的运动方程为

$$\ddot{\theta}+\frac{l}{g}\sin\theta=0,$$

它的位形空间是图 1.2.1(b) 所示的圆周, 即拓扑空间 S^1. 同样, 如果我们考虑图 1.2.2(a) 所示的平面复摆 $0\leqslant\theta\leqslant 360°$, $0\leqslant\varphi\leqslant 360°$, 它的位形空间可用图 1.2.2(b) 所示的环面来表示, 即拓扑空间 T^2.

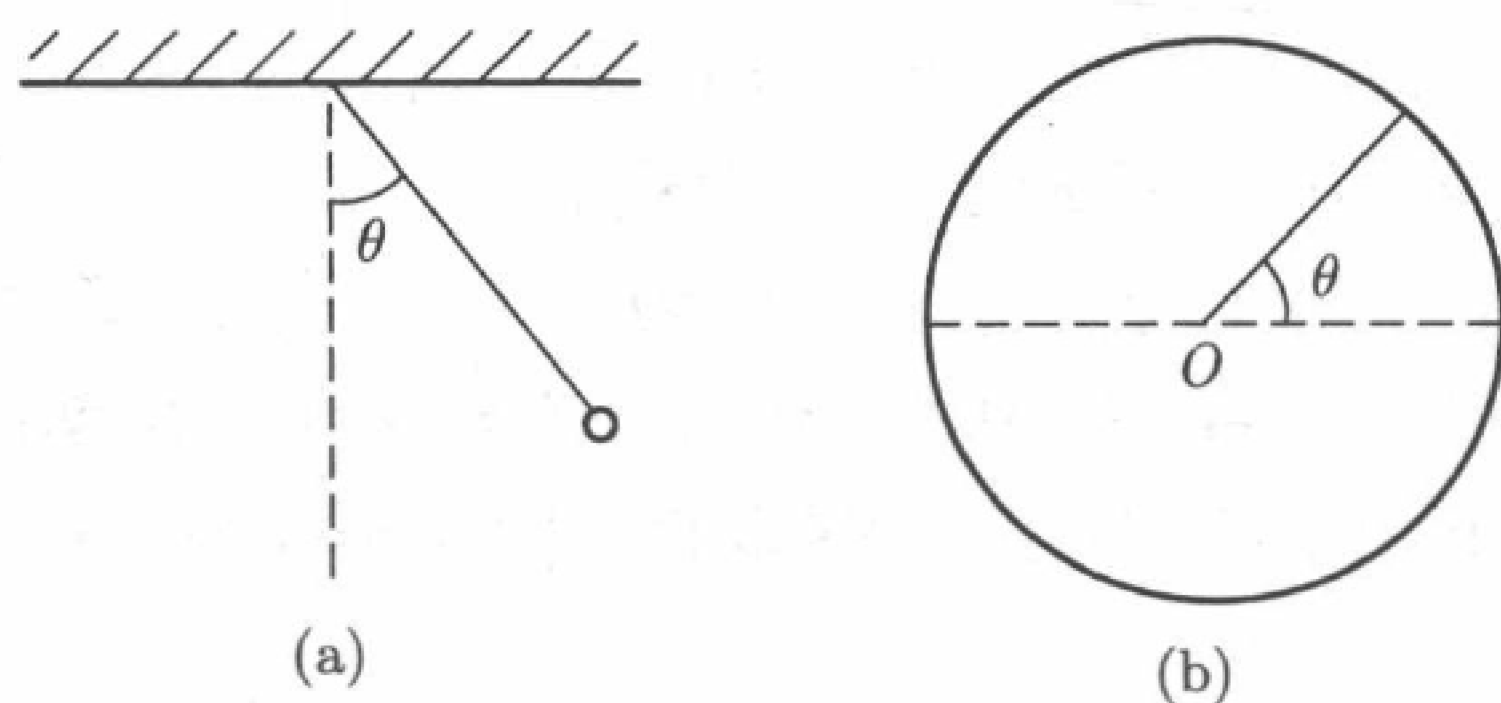


图 1.2.1 (a) 平面单摆; (b) 圆周

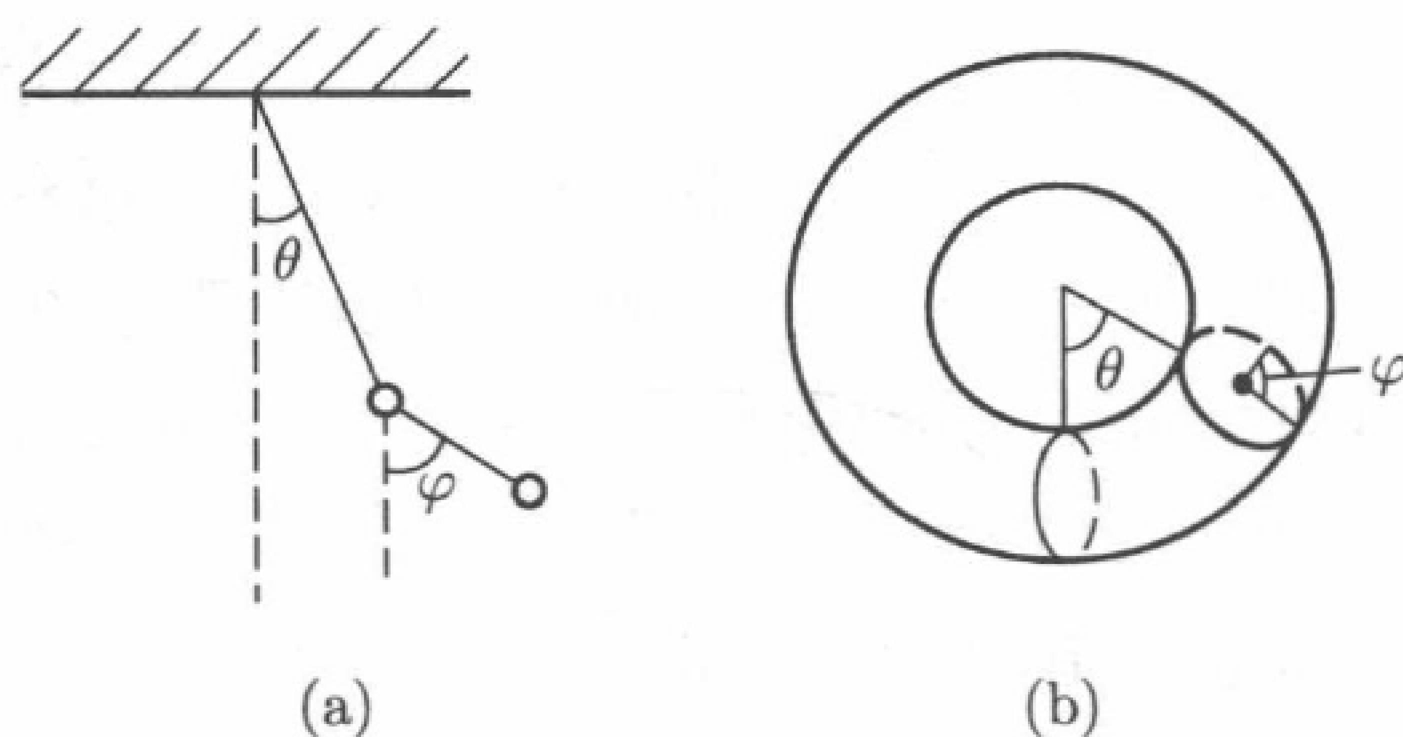


图 1.2.2 (a) 平面复摆; (b) 环面

复摆的运动方程为

$$\ddot{\theta} = -\frac{g(2m_1+m_2)\sin\theta - m_2 g\sin(\theta-\varphi)m_2(\dot{\phi}^2L_2+\dot{\theta}^2L_1\cos(\theta-\varphi))}{L_1(2m_1+m_2-m_2\cos(\theta-\varphi))},$$

$$\ddot{\varphi} = \frac{g(m_1+m_2)\cos\theta+2\sin(\theta-\varphi)(\dot{\theta}^2L_1(m_1+m_2)+\dot{\varphi}^2L_2m_2\cos(\theta-\varphi))}{L_2(2m_1+m_2-m_2\cos(\theta-\varphi))}.$$

S^1 和 T^2 都有局部坐标, 也就是说它们局部同胚于 Euclid 空间, 这就引出了流形的概念. 简单地讲, 拓扑空间加上局部坐标就构成流形, 也就是物理学家所说的位形空间.

已知 R 是拓扑空间, 若能在 R 上取定原点 O, 并规定方向和单位长度, 那么每一点都对应一个坐标, 这样就建立了一个坐标系. 在一般的拓扑空间上是否能建立整体坐标系呢? 一般来说是不能的. 下面以地图册的例子来形象地说明什么是流形. 我们出行的时候, 会用平面地图来指示方位. 如果将整个地球的各个地区的地图合订成一本地图集, 那么在观看各个地区的地图后, 我们就可以在脑海中 "拼接" 出整个地球的景貌. 为了能让阅读者顺利地从一张地图接到下一张, 相邻的地图之间会有重叠的部分, 以便我们在脑海里 "黏合" 两张图. 类似地, 在数学中, 也可以用一系列 "地图" (称为坐标图或坐标卡) 组成的 "地图集" (称为图册) 来描述一个流形. 而 "地图" 之间重叠的部分在不同的地图里如何变换, 则描述了不同 "地图" 之间的关系.

如图 1.2.3(a) 所示, 将点 P 与南极 S 相连, 设连线 PS 或其延长线与赤道平面 xy 的交点 Q 的坐标为 (u,v), 于是可定义 (u,v) 作为点 P 的坐标. 这样就得到一个局部坐标系 S_1, 其定义域为 $2z>-1$, 即在除去南极点 S 的球面上点的坐标都可以用 (u,v) 表示. 类似地, 可以通过北极点 N 按图 1.2.3(b) 的方法, 在除去北极点 N 的球面上建立一个局部坐标系 S_2, 它的局部坐标为 $(\overline{u},\overline{v})$, 定义域为 $2z<1$.

对于二维球面 S^2 上的点 $P\in S_1\cap S_2$, 它有两种坐标, 不难求出它们之间的变换关系

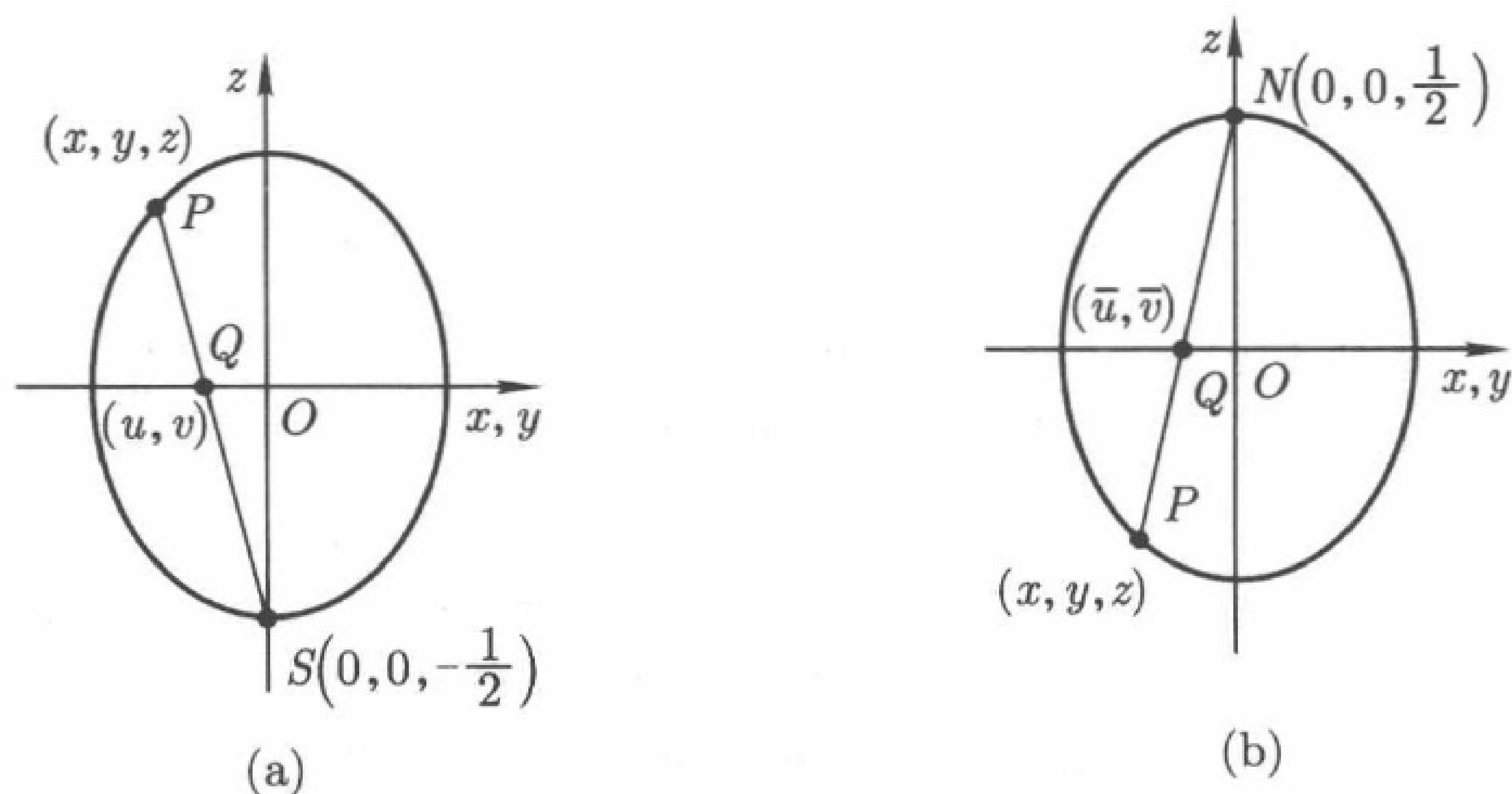


图 1.2.3　位形空间二维球面 S^2

$$\overline{u}=\frac{u}{u^2+v^2},\quad u=\frac{\overline{u}}{\overline{u}^2+\overline{v}^2},$$

$$\overline{v}=\frac{v}{u^2+v^2},\quad v=\frac{\overline{v}}{\overline{u}^2+\overline{v}^2}.$$

变换属于 C^∞ 函数类.

下面我们通过图 1.2.4 来简单示范流形上函数可微性的定义.

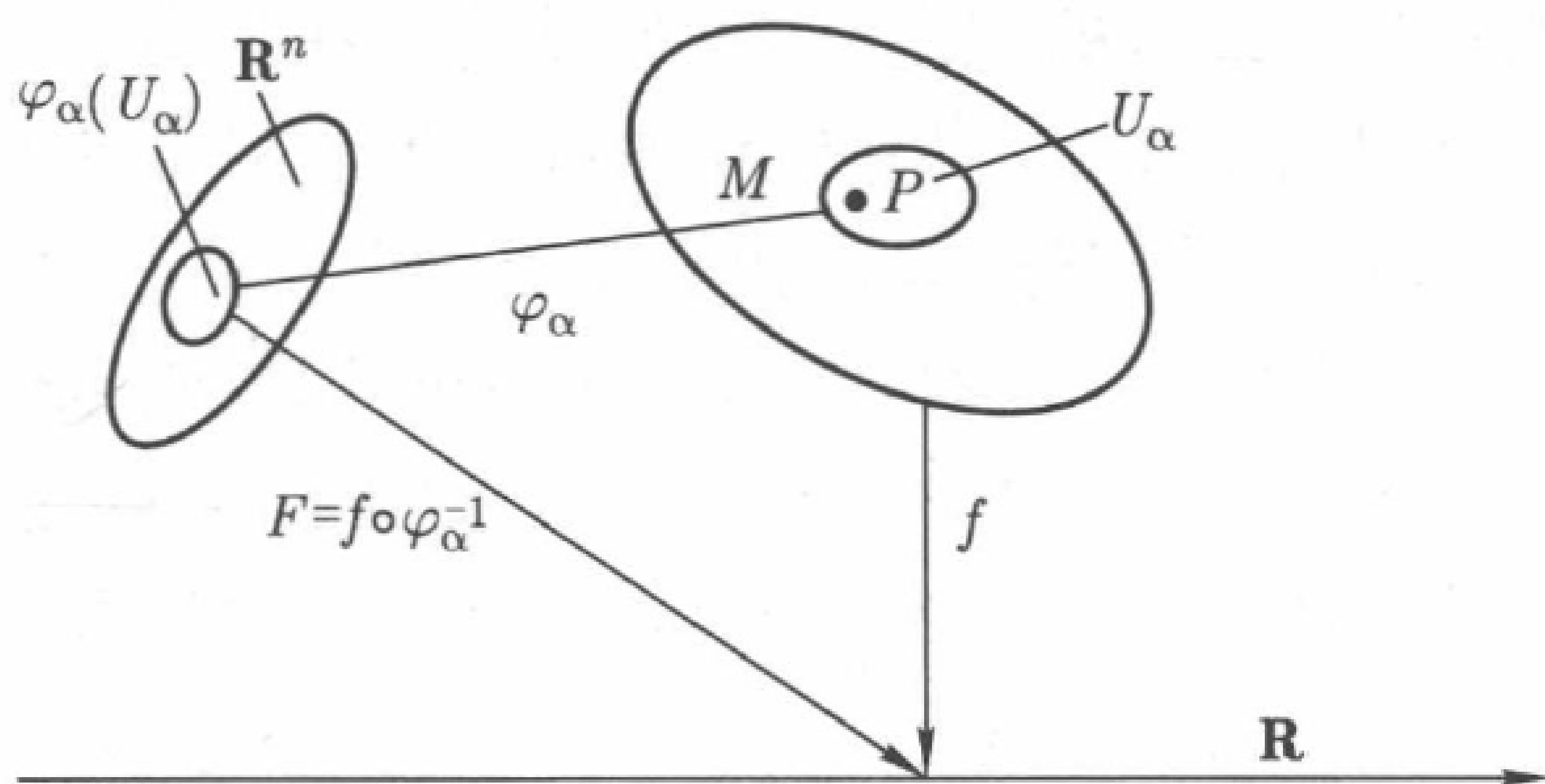


图 1.2.4　函数的可微性

设 M 是一个流形, $\mathbf{R}^n$ 是 n 维 Euclid 空间, f 是定义在 M 上的函数:

$$f:\ M\to\mathbf{R},$$

$$M\ni P\mapsto f(P)\in\mathbf{R},$$

又设 $(U_\alpha, \varphi_\alpha)$ 为含点 P 的一个坐标卡, 则 $F = f \circ \varphi_\alpha^{-1}$ 为定义在 $\mathbf{R}^n$ 开集上的实函数 (图 1.2.4).

我们称 φ_α 为 U_α 的**局部坐标系**, 称 U_α 为 φ_α 的**坐标邻域**, 称 $(U_\alpha, \varphi_\alpha)$ 为**坐标卡**. 所有的坐标卡组成 M 上的一个光滑图册. 下面通过坐标卡 $(U_\alpha, \varphi_\alpha)$ 来定义 C^∞ 流形 M.

定义 1.2.1 设 M 是一个有可数基的 Hausdorff 空间, 给定一个图册 $\mathcal{A} = \{(U_\alpha, \varphi_\alpha)\}$, 满足:

(1) $\{U_\alpha\}$ 是 M 的一个开覆盖;

(2) 对 $\forall \alpha$, 映射 $\varphi_\alpha : U_\alpha \to \varphi_\alpha(U_\alpha)$ 是一个从 M 的开集 U_α 到局部 Euclid 空间 $\mathbf{R}^n$ 上的同胚映射, 如图 1.2.5 所示;

(3) $\mathcal{A}$ 中的任何两个图册 $(U_\alpha, \varphi_\alpha)$ 和 (U_β, φ_β) 是 C^∞ 兼容的 (表示 $U_\alpha \cap U_\beta = \varnothing$, 或者 $U_\alpha \cap U_\beta \neq \varnothing$, 且 $\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$ 是一个微分同胚),

则称 M 为一个 n 维**光滑流形**, 称 $\mathcal{A}$ 为 M 的一个**光滑图册**. 如果 $\mathcal{A}$ 包含所有与它相容的局部坐标系, 则称 $\mathcal{A}$ 为 M 的**最大光滑图册**, 此时称 $(M, \mathcal{A})$ 为 n 维**微分流形** (简称为**流形**), 称 $\mathcal{A}$ 为 M 的**微分结构**.

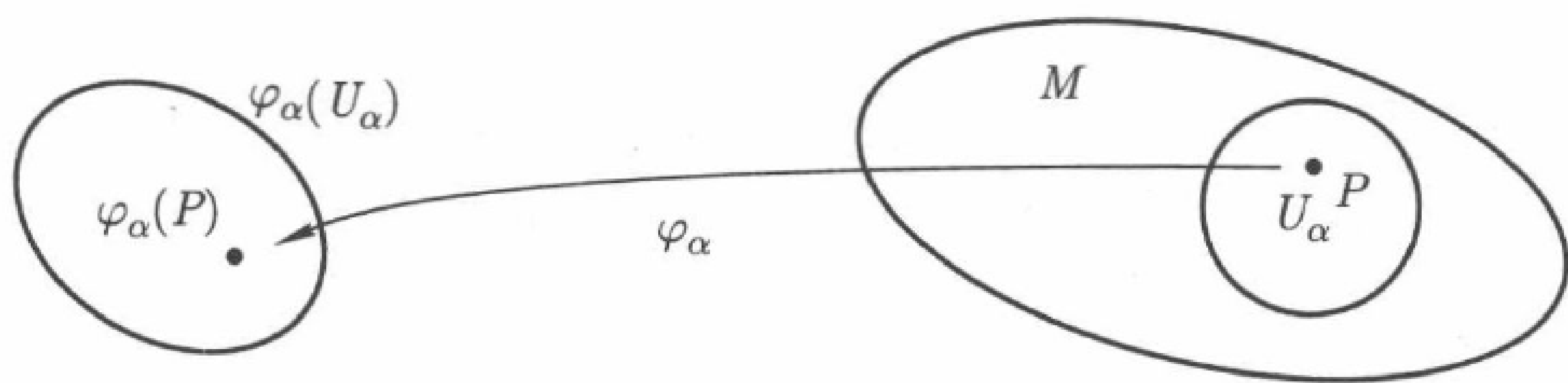


图 1.2.5 M 的开集同胚映射 $\varphi_\alpha : U_\alpha \to \varphi_\alpha(U_\alpha)$($\mathbf{R}^n$ 中的开集)

定义 1.2.2 设 M 和 N 分别是 m 维和 n 维微分流形, 连续映射 $f : M \to N$ 满足对于 $P \in M$ 和 $f(P) \in N$ 的坐标卡 $(U, \varphi), (V, \psi)$, 有 $f(U) \subset V$, 且局部表示 $\widehat{f} = \psi \circ f \circ \varphi^{-1} : \varphi(U) \to \psi(V)$ 是 C^k 可微的, 则称 f 在 P 处是 C^k **可微**的; 如果 f 在每个点 $P \in M$ 处都是 C^k 可微的, 则称 f 是 C^k **可微**的, 或称 f 是 C^k **映射** (图 1.2.6).

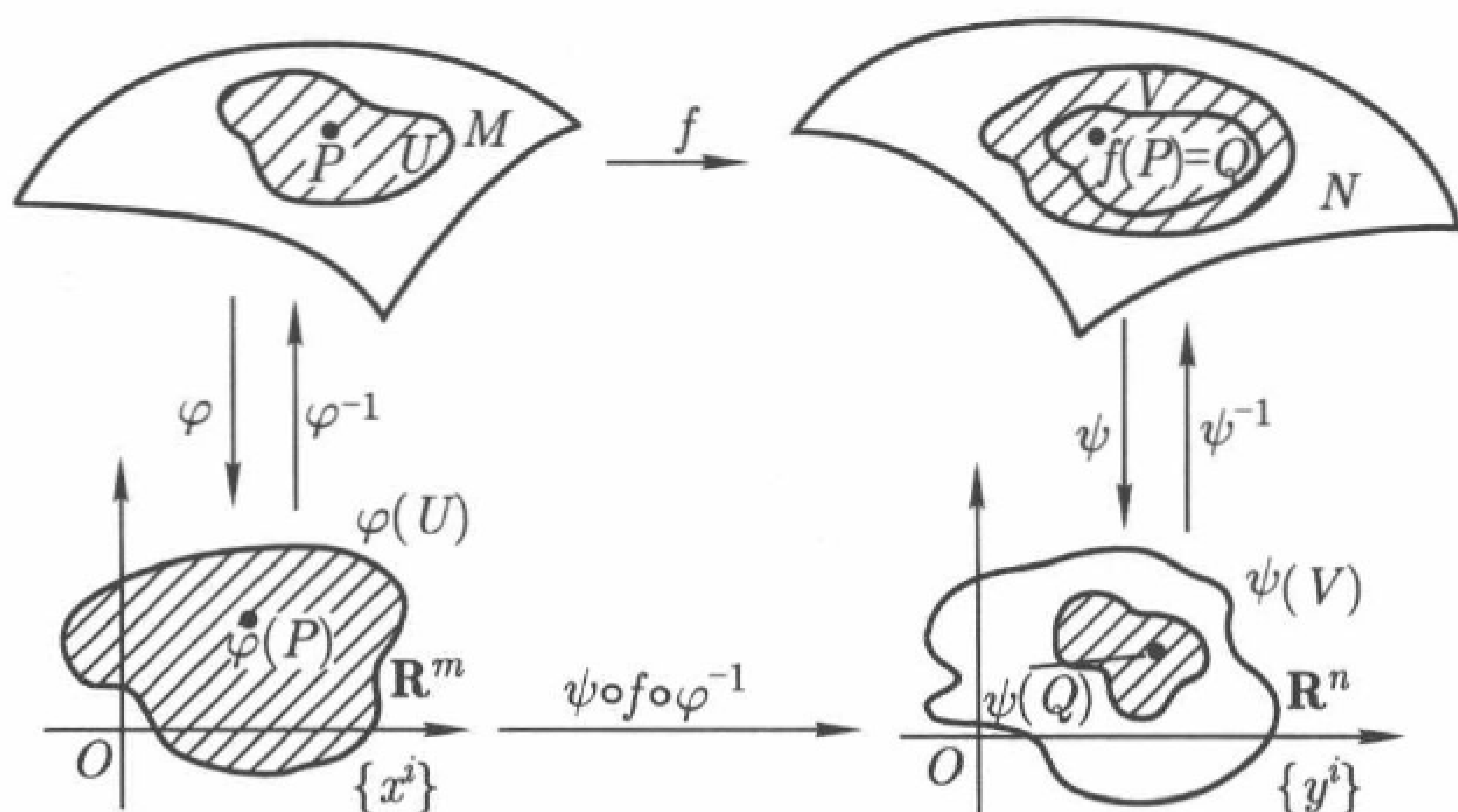


图 1.2.6 可微映射

1.2.2 切空间和余切空间

令 M 是一个 n 维微分流形, $\mathcal{F}_{\boldsymbol{p}}$ 是所有在点 $\boldsymbol{p}=(p_1,\cdots,p_n)\in M$ 附近有定义而且在 $\boldsymbol{p}$ 处可微的函数构成的空间. 微分流形 M 在 $\boldsymbol{p}\in M$ 处的切向量 $X_{\boldsymbol{p}}$ 是指映射

$$X_{\boldsymbol{p}}:\mathcal{F}_{\boldsymbol{p}}\to\mathbf{R},$$

它具有性质:

(1) $X_{\boldsymbol{p}}(f)=X_{\boldsymbol{p}}(g)$, 当 $f,g\in\mathcal{F}_{\boldsymbol{p}}$ 在 $\boldsymbol{p}$ 的某个邻域相等时;

(2) $X_{\boldsymbol{p}}(\alpha f+\beta g)=\alpha X_{\boldsymbol{p}}(f)+\beta X_{\boldsymbol{p}}(g),\forall f,g\in\mathcal{F}_{\boldsymbol{p}},\forall\alpha,\beta\in\mathbf{R}$;

(3) $X_{\boldsymbol{p}}(fg)=f(\boldsymbol{p})X_{\boldsymbol{p}}(g)+g(\boldsymbol{p})X_{\boldsymbol{p}}(f),\forall f,g\in\mathcal{F}_{\boldsymbol{p}}$ (相当于求导运算的莱布尼茨法则).

下面给出切向量的一般定义:

定义 1.2.3 假设一个映射 $\nu:\mathcal{F}_{\boldsymbol{p}}\to\mathbf{R}$ 有下列性质:

(1) 线性性:

$$\nu(\alpha f+\beta g)=\alpha\nu(f)+\beta\nu(g),\quad \alpha,\beta\in\mathbf{R};$$

(2) 莱布尼茨法则:

$$\nu(fg)=f(\boldsymbol{p})\nu(g)+g(\boldsymbol{p})\nu(f),$$

则称 ν 是 M 在点 $\boldsymbol{p}$ 处的一个**切向量** (图 1.2.7).

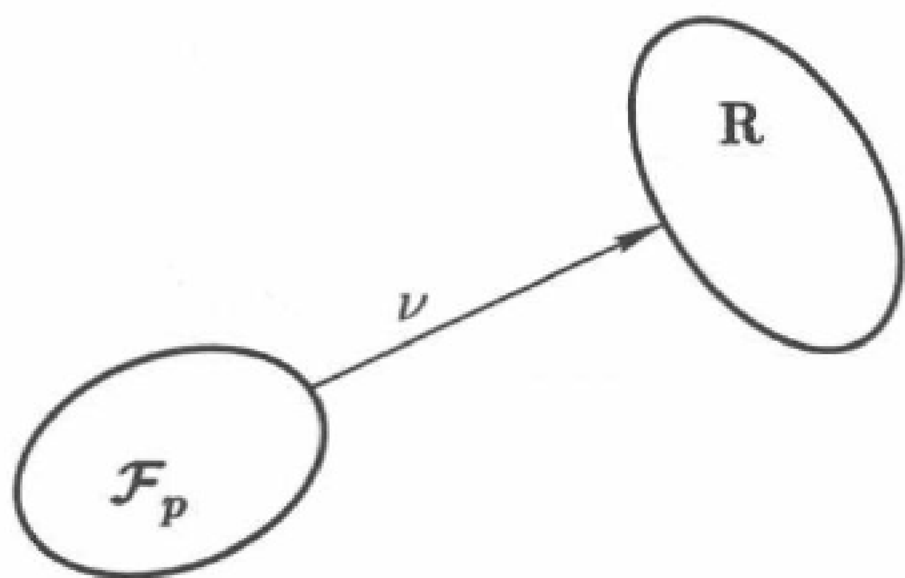


图 1.2.7 点 $\boldsymbol{p}$ 处的切向量

记 $T_{\boldsymbol{p}}M$ ={微分流形 M 在点 $\boldsymbol{p} \in M$ 处的全体切向量}. 假设 $X_{\boldsymbol{p}}$ 和 $Y_{\boldsymbol{p}}$ 是 M 在点 $\boldsymbol{p}$ 处的任意切向量, 定义下面的加法和数乘运算:

$$(X_{\boldsymbol{p}} + Y_{\boldsymbol{p}})(f) = X_{\boldsymbol{p}}(f) + Y_{\boldsymbol{p}}(f),$$
$$(kX_{\boldsymbol{p}})(f) = kX_{\boldsymbol{p}}(f), \quad \forall f \in \mathcal{F}_{\boldsymbol{p}}, \quad \forall k \in \mathbf{R}.$$

易证 $T_{\boldsymbol{p}}M$ 关于上述运算构成向量空间, 我们称其为微分流形 M 在点 $\boldsymbol{p}$ 处的切空间.

定义 1.2.4 由微分流形上一点 $\boldsymbol{p}$ 处的全体切向量构成的向量空间 $T_{\boldsymbol{p}}M$ 称为 M 在点 $\boldsymbol{p}$ 处的**切空间**.

为了后面便于理解切向量的局部坐标运算, 下面我们不加证明地给出切空间基底的有关结论.

定理 1.2.1 假设 $\boldsymbol{x} = (x_1, \cdots, x_n)$ 是微分流形 M 的一个给定的局部坐标系, $\boldsymbol{x}^0 = (x_1^0, \cdots, x_n^0)$ 是点 $\boldsymbol{p}$ 的局部坐标, 则切空间 $T_{\boldsymbol{p}}M$ 是一个 n 维线性空间, 它的基底在给定的坐标系下可以表示为

$$\left.\frac{\partial}{\partial x_1}\right|_{\boldsymbol{x}=\boldsymbol{x}^0}, \quad \cdots, \quad \left.\frac{\partial}{\partial x_n}\right|_{\boldsymbol{x}=\boldsymbol{x}^0},$$

$T_{\boldsymbol{p}}M$ 中的每个切向量 ν 都可以表示为该组基底的一个线性组合:

$$\nu = \alpha_1 \left.\frac{\partial}{\partial x_1}\right|_{\boldsymbol{x}=\boldsymbol{x}^0} + \cdots + \alpha_n \left.\frac{\partial}{\partial x_n}\right|_{\boldsymbol{x}=\boldsymbol{x}^0}, \quad \alpha_1, \cdots, \alpha_n \in \mathbf{R}.$$

定义 1.2.5 设 V 为域 F 上的向量空间, 其对偶空间 V^* 为由 V 到 F 的所有线性函数组成的集合, 即 V 的标量线性变换. V^* 本身是 F 的向量空间并且拥有加法及标量乘法运算. 在张量的语言中, V 的元

素被称为**逆变向量**, 而 V^* 的元素被称为**协变向量**、**合同向量**或 **1-形式**.

定义 1.2.6　$T_{\boldsymbol{p}}M$ 的对偶空间称为**余切空间**, 记为 $T_{\boldsymbol{p}}^*M$. 余切空间的基为

$$\mathrm{d}x_1, \quad \cdots, \quad \mathrm{d}x_n,$$

它是切空间基向量

$$\frac{\partial}{\partial x_1}, \quad \cdots, \quad \frac{\partial}{\partial x_n}$$

的对偶基.

定义 1.2.7　三元组 (TM, M, π) 称为微分流形 M 的**切丛**, 其中 $TM = \bigcup\limits_{\boldsymbol{p}\in M} T_{\boldsymbol{p}}M$, 投影映射 $\pi : TM \to M$ 满足 $\pi(X_{\boldsymbol{p}}) = \boldsymbol{p}$, $\forall X_{\boldsymbol{p}} \in TM$. 通常将切丛 (TM, M, π) 简记为 TM. 对于每个 $\boldsymbol{p} \in M$, $\pi^{-1}(\boldsymbol{p}) = T_{\boldsymbol{p}}M$ 称为切丛 TM 在点 $\boldsymbol{p}$ 处的**纤维**.

定义 1.2.8　设 $f : M \to N$ 是从流形 M 到流形 N 的光滑映射, $f_{*\boldsymbol{x}}$ 表示切空间的诱导映射, $f_{*\boldsymbol{x}} : T_{\boldsymbol{x}}M \to T_{f(\boldsymbol{x})}N$. 若点 $\boldsymbol{x}$ 在 M 上变化, 则 $f_{*\boldsymbol{x}}$ 定义了一个从切丛 TM 到切丛 TN 的**切映射**:

$$f_* : TM \to TN,$$

$$f_*|_{T_{\boldsymbol{x}}M} = f_{*\boldsymbol{x}}.$$

特别地, 若 $N = M = \mathbf{R}^n$, 则

$$f_{*\boldsymbol{x}} : T_{\boldsymbol{x}}\mathbf{R}^n \to T_{f(\boldsymbol{x})}\mathbf{R}^n,$$

$$f_{*\boldsymbol{x}}\xi = \frac{\partial f}{\partial \boldsymbol{x}}\xi.$$

切映射的几何意义如图 1.2.8 所示.

定义 1.2.9　设 M 是一个微分流形, 如果映射 $X : M \to TM$ 满足 $\pi \circ X$ 是 M 上的恒等映射, 则称 X 为 M 的一个**向量场**, 也称 X 为切丛 TM 的一个**截面**; 如果 X 还是光滑映射, 则称它为 M 的一个**光滑向量场**.

微分流形 M 的所有光滑向量场构成的集合记为 $\chi(M)$.

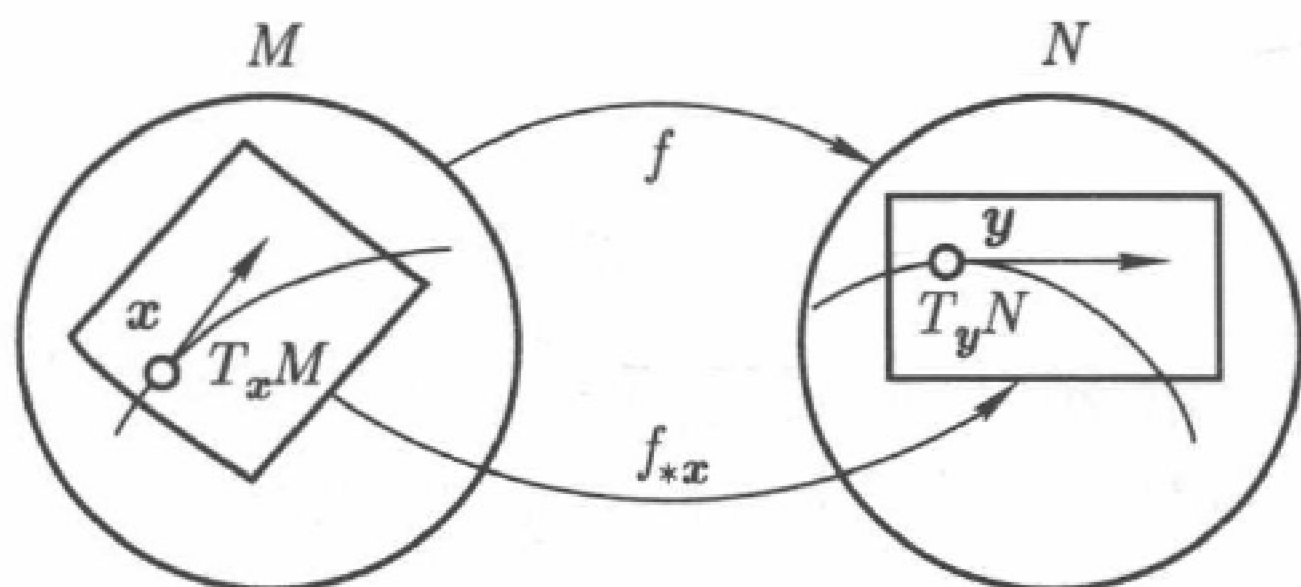


图 1.2.8 切映射的几何意义

特别地, 当 $X \in \chi(\mathbf{R}^n)$ 时, X 的坐标表示为

$$X = \sum_{i=1}^{n} a_i(\boldsymbol{x}) \frac{\partial}{\partial x_i}.$$

假设 $f \in C^1(\mathbf{R}^n)$, 那么有

$$Xf = \sum_{i=1}^{n} a_i(\boldsymbol{x}) \frac{\partial f}{\partial x_i}.$$

定义 1.2.10 设 X 是微分流形 M 的一个光滑向量场, 光滑映射 $c: [0,T] \to M$ 过点 $c(0) = \boldsymbol{x}^0 \in M$, 且满足

$$c_{*t}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right) = X_{c(t)}, \quad \forall t \in [0,T], \tag{1.2.1}$$

其中 c_{*t} 是映射 c 的切映射, 则称 $c(t)$ 为向量场 X 的过点 $\boldsymbol{x}^0$ 的**解曲线**. 此时, 存在群作用映射 $\varphi: c[0,T] \times [0,T] \to c[0,T]$, 满足

$$\varphi(\boldsymbol{x}^0, 0) = \boldsymbol{x}^0,$$
$$\varphi(\varphi(\boldsymbol{x}, t_1), t_2) = \varphi(\boldsymbol{x}, t_1 + t_2), \quad t_1, t_2 \in [0,T],$$

称 $\varphi(\boldsymbol{x}, t)$ 为 X 的**流**, 简记为 φ^t. 特别地, 当 X 是 Hamilton 向量场时, 称 X 的流为**相流**.

§1.3 代数的预备知识

首先简单回忆一下解析几何中向量的点乘运算和叉乘运算. 两个向量的点乘也叫做向量的数量积或内积. 顾名思义, 点乘运算的结果是

一个数. 向量点乘的运算公式为

$$\boldsymbol{a}\cdot\boldsymbol{b}=|\boldsymbol{a}||\boldsymbol{b}|\cos\langle\boldsymbol{a},\boldsymbol{b}\rangle,$$

其中 $\langle\boldsymbol{a},\boldsymbol{b}\rangle$ 表示向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 的几何夹角. 在物理学中, 已知物体的位移 $\boldsymbol{s}$ 以及作用在物体上的力 $\boldsymbol{F}$, 则力 $\boldsymbol{F}$ 所做的功 $W=\boldsymbol{F}\cdot\boldsymbol{s}$. 两个向量的叉乘也叫做向量的向量积或外积. 顾名思义, 叉乘运算的结果是一个向量. 向量叉乘 $\boldsymbol{c}=\boldsymbol{a}\times\boldsymbol{b}$ 的运算公式为

$$|\boldsymbol{c}|=|\boldsymbol{a}\times\boldsymbol{b}|=|\boldsymbol{a}||\boldsymbol{b}|\sin\langle\boldsymbol{a},\boldsymbol{b}\rangle,$$

向量 $\boldsymbol{c}$ 的方向与 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 所在的平面垂直, 且方向要用 "右手法则" 来判断 (用右手的四指先表示向量 $\boldsymbol{a}$ 的方向, 然后手指朝着手心的方向摆动到向量 $\boldsymbol{b}$ 的方向, 大拇指所指的方向就是向量 $\boldsymbol{c}$ 的方向) . 显然, 向量的叉乘不遵守交换律, 因为 $\boldsymbol{a}\times\boldsymbol{b}=-\boldsymbol{b}\times\boldsymbol{a}$. 在物理学中, 已知力与力臂求力矩, 就是利用向量的叉乘. 将向量用直角坐标表示 (三维向量) , 若向量 $\boldsymbol{a}=(a_1,a_2,a_3)$, $\boldsymbol{b}=(b_1,b_2,b_3)$, 则点乘和叉乘的坐标计算公式分别为

$$\boldsymbol{a}\cdot\boldsymbol{b}=a_1b_1+a_2b_2+a_3b_3,$$

$$\boldsymbol{a}\times\boldsymbol{b}=\begin{vmatrix}\boldsymbol{i}&\boldsymbol{j}&\boldsymbol{k}\\a_1&a_2&a_3\\b_1&b_2&b_3\end{vmatrix}=(a_2b_3-b_2a_3,a_3b_1-a_1b_3,a_1b_2-b_1a_2),$$

其中 $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$ 为三维空间直角坐标系的单位坐标向量.

另外, 还有向量 $\boldsymbol{a}=(a_1,a_2,a_3),\boldsymbol{b}=(b_1,b_2,b_3),\boldsymbol{c}=(c_1,c_2,c_3)$ 的三重积, 也称为混合积:

$$\boldsymbol{a}\times\boldsymbol{b}\cdot\boldsymbol{c}=\begin{vmatrix}a_1&b_1&c_1\\a_2&b_2&c_2\\a_3&b_3&c_3\end{vmatrix}.$$

在三维空间中叉乘和三重积都有对应的几何意义. 叉乘 $\boldsymbol{a}\times\boldsymbol{b}$ 的长度 $|\boldsymbol{a}\times\boldsymbol{b}|$ 等于以向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 为边的平行四边形的面积, 向量 $\boldsymbol{a},\boldsymbol{b}$ 和 $\boldsymbol{c}$ 的三重积 $\boldsymbol{a}\times\boldsymbol{b}\cdot\boldsymbol{c}$ 的几何意义是以 $\boldsymbol{a}$, $\boldsymbol{b}$ 和 $\boldsymbol{c}$ 为边的平行六面体的带有正、负号的体积 (定向体积).

1.3.1 外积

一般的向量外积或者称 $\wedge$ 积, 是向量叉乘向更高维空间的推广. 所以, 和向量叉乘一样, 一般的向量外积在 Euclid 几何中就是用来研究面积、体积以及高维几何空间中类似面积和体积等的几何量的. 和向量叉乘一样, 一般的向量外积也具有反对称性, 即

$$\boldsymbol{u} \wedge \boldsymbol{v} = -\boldsymbol{v} \wedge \boldsymbol{u}$$

对所有向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 都成立.

几何背景: 假设有下面三个三维向量

$$\begin{aligned}
\boldsymbol{a}_1 &= a_{11}\boldsymbol{i} + a_{12}\boldsymbol{j} + a_{13}\boldsymbol{k}, \\
\boldsymbol{a}_2 &= a_{21}\boldsymbol{i} + a_{22}\boldsymbol{j} + a_{23}\boldsymbol{k}, \\
\boldsymbol{a}_3 &= a_{31}\boldsymbol{i} + a_{32}\boldsymbol{j} + a_{33}\boldsymbol{k},
\end{aligned}$$

在 $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ 之间引进运算 $\wedge$ 如下:

$$\boldsymbol{a}_1 \wedge \boldsymbol{a}_2 \wedge \boldsymbol{a}_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}. \quad \text{(定向体积)}$$

该运算具有下面的性质:

(1) **多重线性性**:

$$\boldsymbol{a}_1 \wedge (\beta\boldsymbol{b} + \gamma\boldsymbol{c}) \wedge \boldsymbol{a}_3 = \beta(\boldsymbol{a}_1 \wedge \boldsymbol{b} \wedge \boldsymbol{a}_3) + \gamma(\boldsymbol{a}_1 \wedge \boldsymbol{c} \wedge \boldsymbol{a}_3), \quad \beta, \gamma \in \mathbf{R};$$

(2) **反交换性**:

$$\boldsymbol{a}_1 \wedge \boldsymbol{a}_2 \wedge \boldsymbol{a}_3 = -\boldsymbol{a}_2 \wedge \boldsymbol{a}_1 \wedge \boldsymbol{a}_3 = \boldsymbol{a}_2 \wedge \boldsymbol{a}_3 \wedge \boldsymbol{a}_1.$$

若 $\boldsymbol{a}_1 = \boldsymbol{a}_2$, 则

$$\boldsymbol{a}_1 \wedge \boldsymbol{a}_2 \wedge \boldsymbol{a}_3 = 0.$$

将 $\wedge$ 运算所满足的多重线性性和反交换性加以抽象, 就是一般的向量外积的概念.

假设 e_1, e_2, e_3 是 $\mathbf{R}^3$ 中的一组基 (不一定是正交基), 向量 a_1,a_2,a_3 可以表示为

$$a_1 = a_{11}e_1 + a_{12}e_2 + a_{13}e_3,$$
$$a_2 = a_{21}e_1 + a_{22}e_2 + a_{23}e_3,$$
$$a_3 = a_{31}e_1 + a_{32}e_2 + a_{33}e_3.$$

例 1.3.1　三个向量 a_1, a_2, a_3 的外积为

$$a_1 \wedge a_2 \wedge a_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} e_1 \wedge e_2 \wedge e_3.$$

例 1.3.2　两个向量 a_1, a_2 的外积为

$$a_1 \wedge a_2 = \underset{A_{12}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} e_1 \wedge e_2 + \underset{A_{23}}{\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}} e_2 \wedge e_3 + \underset{A_{31}}{\begin{vmatrix} a_{13} & a_{11} \\ a_{23} & a_{21} \end{vmatrix}} e_3 \wedge e_1. \tag{1.3.1}$$

在 $\mathbf{R}^3$ 中, 坐标平面 e_1e_2 是由坐标向量 e_1 和 e_2 确定的坐标平面, 其他两个坐标平面 e_2e_3 和 e_3e_1 类似. 假设向量 a_1 和 a_2 张成的平行四边形是 A:

$$A = \{x \in \mathbf{R}^3 | x = \alpha a_1 + \beta a_2,\ 0 \leqslant \alpha, \beta \leqslant 1\}.$$

$a_1 \wedge a_2$ 的几何意义如图 1.3.1 所示, 就是由 a_1 和 a_2 在坐标平面上的投影所张成的平行四边形. 所以 (1.3.1) 式中 A_{12}, A_{23} 和 A_{31} 表示的项分别等价于平行四边形 A 在坐标平面 e_1e_2, e_2e_3 和 e_3e_1 上的投影, 即

A 投影到坐标平面 e_1e_2 是 A_{12},

A 投影到坐标平面 e_2e_3 是 A_{23},

A 投影到坐标平面 e_3e_1 是 A_{31}.

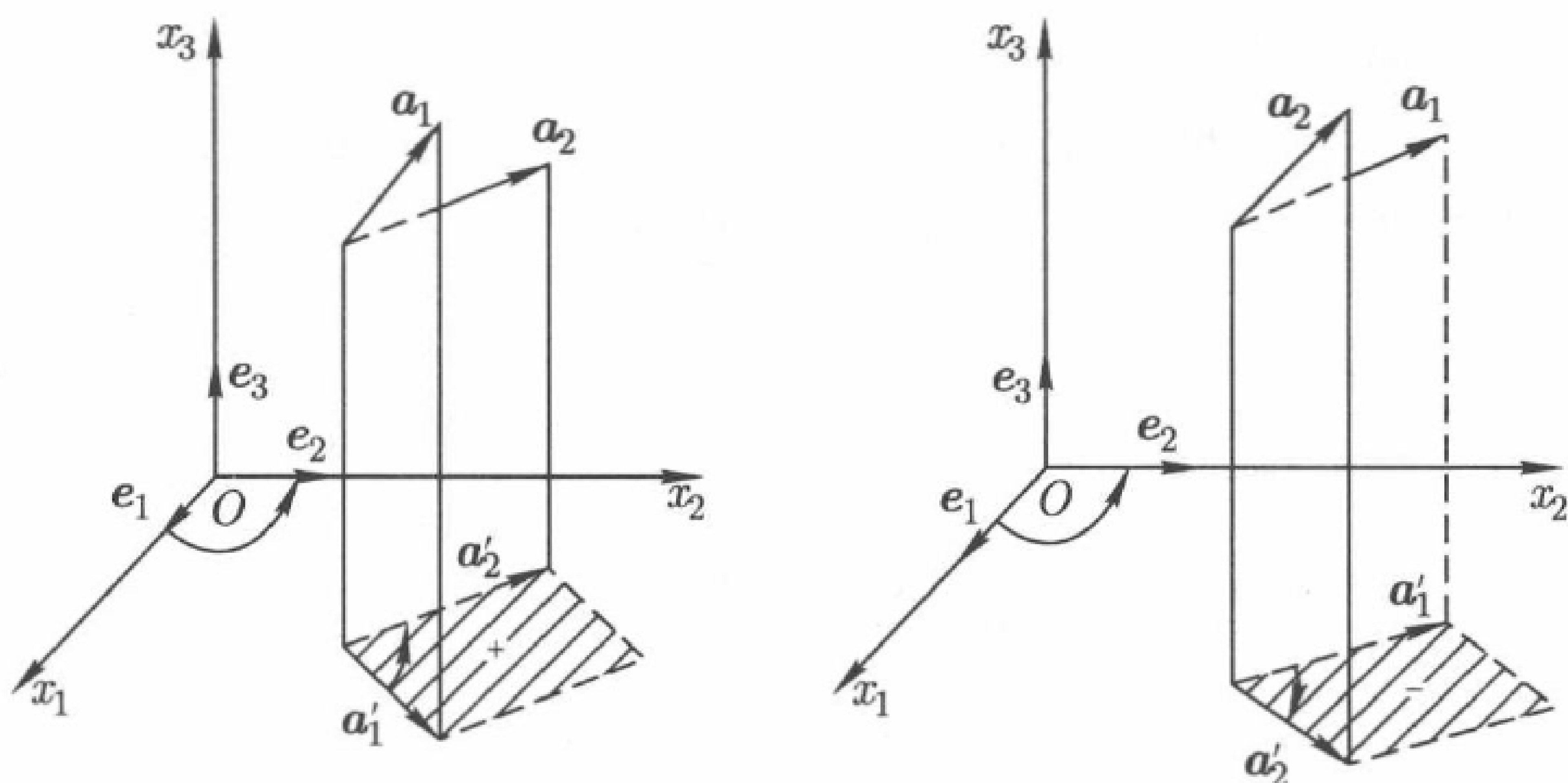


图 1.3.1 外积算子 ∧ 的几何意义

1.3.2 外形式

设 $\mathbf{R}$ 是实数域, V 是 n 维实线性空间, V 中的元素用 $\boldsymbol{u}, \boldsymbol{v}$ 等表示, V 有一组基 $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$. 对 $k = 0, 1, \cdots, n$, 构造实数域上新的线性空间即所谓的 k-**形式线性空间** Λ^k 如下:

当 $k = 0, 1$ 时,

$$\Lambda^0 = \mathbf{R}, \quad \Lambda^1 = V.$$

当 $k = 2$ 时, 作如下定义:

定义 1.3.1 Λ^2 由所有形如

$$\sum_i \alpha_i(\boldsymbol{u}_i \wedge \boldsymbol{v}_i),$$

的元素组成, 其中 $\alpha_i \in \mathbf{R}, \boldsymbol{u}_i, \boldsymbol{v}_i \in V$, 并对 $\forall \boldsymbol{u}, \boldsymbol{v} \in V$, 要求 $\wedge$ 满足下列条件:

(1) **多重线性性**:

$$(\alpha_1\boldsymbol{u}_1 + \alpha_2\boldsymbol{u}_2) \wedge \boldsymbol{v} = \alpha_1(\boldsymbol{u}_1 \wedge \boldsymbol{v}) + \alpha_2(\boldsymbol{u}_2 \wedge \boldsymbol{v}),$$
$$\boldsymbol{u} \wedge (\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2) = \alpha_1(\boldsymbol{u} \wedge \boldsymbol{v}_1) + \alpha_2(\boldsymbol{u} \wedge \boldsymbol{v}_2);$$

(2) **反交换性**:

$$\boldsymbol{u} \wedge \boldsymbol{v} = -\boldsymbol{v} \wedge \boldsymbol{u};$$

(3) $\boldsymbol{u}\wedge\boldsymbol{u}=0$.

$\boldsymbol{u}\wedge\boldsymbol{v}$ 就是 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的外积. Λ^2 的元素称为 **2-形式**.

在给定基 $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$ 时, 我们有

$$\boldsymbol{u}=\sum_{i=1}^n \alpha_i\boldsymbol{e}_i,\quad \boldsymbol{v}=\sum_{i=1}^n \beta_i\boldsymbol{e}_i,\quad \alpha_i,\beta_i\in\mathbf{R},\ i=1,\cdots,n,$$

则根据 $\wedge$ 的多重线性性和反交换性得到

$$\begin{aligned}\boldsymbol{u}\wedge\boldsymbol{v}&=\left(\sum_{i=1}^n \alpha_i\boldsymbol{e}_i\right)\wedge\left(\sum_{i=1}^n \beta_i\boldsymbol{e}_i\right)\\&=\sum_{i,j}\alpha_i\beta_j(\boldsymbol{e}_i\wedge\boldsymbol{e}_j),\end{aligned}$$

因为 $\boldsymbol{e}_i\wedge\boldsymbol{e}_i=0,\boldsymbol{e}_i\wedge\boldsymbol{e}_j=-\boldsymbol{e}_j\wedge\boldsymbol{e}_i$, 所以

$$\boldsymbol{u}\wedge\boldsymbol{v}=\sum_{i<j}(\alpha_i\beta_j-\alpha_j\beta_i)\boldsymbol{e}_i\wedge\boldsymbol{e}_j.$$

由此可见, 线性空间 Λ^2 的元素是特殊形式 $\boldsymbol{e}_i\wedge\boldsymbol{e}_j$ $(1\leqslant i<j\leqslant n)$ 的线性组合. 因此 $\{\boldsymbol{e}_i\wedge\boldsymbol{e}_j\}_{1\leqslant i<j\leqslant n}$ 是 Λ^2 的一组基. Λ^2 的维数为 $\mathrm{C}_n^2=\dfrac{n(n-1)}{2}$, 其元素的一般形式为

$$\sum_{1\leqslant i<j\leqslant n}\alpha_{ij}\boldsymbol{e}_i\wedge\boldsymbol{e}_j,\qquad \alpha_{ij}\in\mathbf{R}.$$

当 $3\leqslant k\leqslant n$ 时, 可以类似于 Λ^2 定义 Λ^k. Λ^k 的元素称为 k-**形式**. 显然, k-形式是多重线性形式. k-形式也称为**外形式**.

1.3.3　微分形式

假设 M 是一个微分流形, TM 是其切丛.

设 $M\ni\boldsymbol{x}\mapsto\omega(\boldsymbol{x})\in\Lambda^k(T_{\boldsymbol{x}}^*M)$, $T_{\boldsymbol{x}}^*M$ 是所有 M 上的光滑 1-形式组成的集合. M 上的微分 k-形式, 在代数上可以视做 M 的余切空间上的反对称多重线性形式. 若 $\boldsymbol{x}$ 有局部坐标 $(x_1,\cdots,x_n)$, 以 $\mathrm{d}x_1,\cdots,\mathrm{d}x_n$ 为基的线性空间称为**微分空间**, 该空间的元素是微分 1-形式:

$$\omega=a_1(\boldsymbol{x})\mathrm{d}x_1+\cdots+a_n(\boldsymbol{x})\mathrm{d}x_n.\tag{1.3.2}$$

在微分空间上, 利用上一节所讲的方法可构造微分 2-形式

$$\sum_{1\leqslant i,j\leqslant n} a_{ij}(\boldsymbol{x})\mathrm{d}x_i \wedge \mathrm{d}x_j \tag{1.3.3}$$

以及微分 k-形式

$$\omega(\boldsymbol{x}) = \sum_{i_1<\cdots<i_k} a_{i_1\cdots i_k}(\boldsymbol{x})\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}, \tag{1.3.4}$$

其中 $a_{i_1\cdots i_k}(\boldsymbol{x}) \in \mathcal{F}(\mathbf{R}^n)$ 是 $\mathbf{R}^n$ 上的光滑函数, 可以看做微分 0-形式.

对于 $f \in \mathcal{F}(\mathbf{R}^n)$, 显然有

$$\mathrm{d}f = \frac{\partial f}{\partial x_1}\mathrm{d}x_1 + \cdots + \frac{\partial f}{\partial x_n}\mathrm{d}x_n.$$

性质 1.3.1 (微分形式的性质) 假设 ω 是微分 k-形式, η 是微分 l-形式, 则有

(1) $\mathrm{d}(\omega+\eta) = \mathrm{d}\omega + \mathrm{d}\eta$;

(2) $\mathrm{d}(\omega\wedge\eta) = \mathrm{d}\omega\wedge\eta + (-1)^k\omega\wedge\mathrm{d}\eta$;

(3) $\mathrm{d}(\mathrm{d}\omega) = \mathrm{d}^2\omega = 0$.

证明 (1) 由微分形式的定义容易得到.

(2) 不妨假设

$$\omega = a(\boldsymbol{x})\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_k}, \quad \eta = b(\boldsymbol{x})\mathrm{d}x_{j_1}\wedge\cdots\wedge\mathrm{d}x_{j_l}.$$

由 $\wedge$ 运算的定义, 我们有

$$\begin{aligned}
\mathrm{d}(\omega\wedge\eta) &= \sum_{i=1}^n \frac{\partial(ab)}{\partial x_i}\mathrm{d}x_i\wedge\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_k}\wedge\mathrm{d}x_{j_1}\wedge\cdots\wedge\mathrm{d}x_{j_l} \\
&= \sum_{i=1}^n \left(b\frac{\partial a}{\partial x_i} + a\frac{\partial b}{\partial x_i}\right)\mathrm{d}x_i\wedge\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_k}\wedge\mathrm{d}x_{j_1}\wedge\cdots\wedge\mathrm{d}x_{j_l} \\
&= \sum_{i=1}^n b\frac{\partial a}{\partial x_i}\mathrm{d}x_i\wedge\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_k}\wedge\mathrm{d}x_{j_1}\wedge\cdots\wedge\mathrm{d}x_{j_l} \\
&\quad + \sum_{i=1}^n a\frac{\partial b}{\partial x_i}(-1)^k\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_k}\wedge\mathrm{d}x_i\wedge\mathrm{d}x_{j_1}\wedge\cdots\wedge\mathrm{d}x_{j_l} \\
&= \mathrm{d}\omega\wedge\eta + (-1)^k\omega\wedge\mathrm{d}\eta.
\end{aligned}$$

$$(3)\ \mathrm{d}\omega=\sum_{i=1}^{n}\frac{\partial a}{\partial x_i}\mathrm{d}x_i\wedge \mathrm{d}x_{i_1}\wedge\cdots\wedge \mathrm{d}x_{i_k},$$

$$\begin{aligned}\mathrm{d}(\mathrm{d}\omega)&=\sum_{i=1}^{n}\mathrm{d}\left(\frac{\partial a}{\partial x_i}\right)\wedge \mathrm{d}x_i\wedge \mathrm{d}x_{i_1}\wedge\cdots\wedge \mathrm{d}x_{i_k},\\&=\left(\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{\partial^2 a}{\partial x_j\partial x_i}\mathrm{d}x_j\wedge \mathrm{d}x_i\right)\wedge \mathrm{d}x_{i_1}\wedge\cdots\wedge \mathrm{d}x_{i_k}\\&=0,\end{aligned}$$

其中

$$\begin{aligned}&\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{\partial^2 a}{\partial x_j\partial x_i}\mathrm{d}x_j\wedge \mathrm{d}x_i\\&=\frac{1}{2}\left(\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{\partial^2 a}{\partial x_j\partial x_i}\mathrm{d}x_j\wedge \mathrm{d}x_i+\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{\partial^2 a}{\partial x_j\partial x_i}\mathrm{d}x_j\wedge \mathrm{d}x_i\right)\\&=\frac{1}{2}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 a}{\partial x_i\partial x_j}\mathrm{d}x_i\wedge \mathrm{d}x_j+\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{\partial^2 a}{\partial x_j\partial x_i}\mathrm{d}x_j\wedge \mathrm{d}x_i\right)\\&=\frac{1}{2}\left(-\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 a}{\partial x_i\partial x_j}\mathrm{d}x_j\wedge \mathrm{d}x_i+\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{\partial^2 a}{\partial x_j\partial x_i}\mathrm{d}x_j\wedge \mathrm{d}x_i\right)\\&=\frac{1}{2}\left(-\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 a}{\partial x_i\partial x_j}+\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{\partial^2 a}{\partial x_j\partial x_i}\right)\mathrm{d}x_j\wedge \mathrm{d}x_i=0.\end{aligned}$$ □

定义 1.3.2 一个双线性的、非退化的反对称 2-形式 $\omega: M\times M\to\mathbf{R}$ 称为**辛形式**, 即 ω 满足

(1) 反对称性: $\omega(\boldsymbol{u},\boldsymbol{v})=-\omega(\boldsymbol{v},\boldsymbol{u}),\forall\boldsymbol{u},\boldsymbol{v}\in M$;

(2) 非退化: 若对 $\forall\boldsymbol{v}\in M$, 有 $\omega(\boldsymbol{u},\boldsymbol{v})=0$, 则必有 $\boldsymbol{u}=\mathbf{0}$.

定义 1.3.3 配备了一个辛微分形式 ω 的微分流形 M 称为**辛流形**, 记为 (M,ω), 其中 ω 称为 M 的**辛结构**.

定义 1.3.4 假设 M 和 N 分别是 m 维和 n 维微分流形, φ 是从 M 到 N 的光滑映射, 则可以把 N 上的微分 k-形式 ω 拉回到 M 上的

微分 k-形式 $\varphi^*\omega$:

$$\varphi^*\omega(\boldsymbol{x})(\xi_1,\cdots,\xi_k)=\omega(\varphi(\boldsymbol{x}))(\varphi_*\xi_1,\cdots,\varphi_*\xi_k),$$

其中 $\xi_1,\cdots,\xi_k\in T_{\boldsymbol{x}}M$. 称 $\varphi^*\omega$ 为 ω 的**拉回映射** (图 1.3.2).

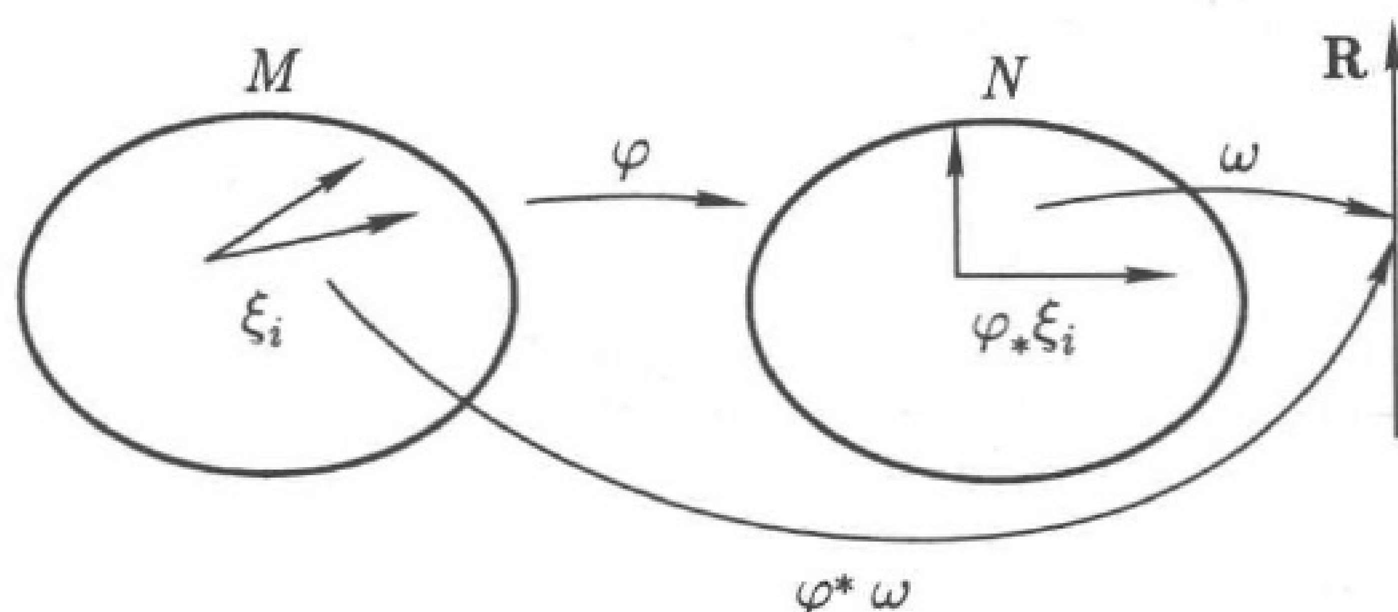


图 1.3.2 微分形式的拉回映射

事实上, 切映射 $\varphi_*:T_{\boldsymbol{x}}M\to T_{\boldsymbol{y}}N$ 会诱导出对偶空间也就是余切空间上的对偶映射

$$\varphi^*:T^*_{\boldsymbol{y}}(N)\to T^*_{\boldsymbol{x}}(M).$$

因为对偶映射的方向与切映射的方向相反, 因此称为拉回映射. 微分流形上的微分形式是切向量场的线性泛函, 微分流形之间的映射 φ 诱导出微分流形上微分形式的拉回映射 φ^*, 即将微分流形 N 上点 $\boldsymbol{y}=\varphi(\boldsymbol{x})$ 处的微分形式拉回到微分流形 M 上点 $\boldsymbol{x}$ 处的微分形式. 例如, 对 0-形式, 即微分流形 N 上的函数 $f(\boldsymbol{y})$, 利用微分流形 M 到 N 的映射 φ, 可将 N 上的函数 f 拉回到 M 上的函数 $g(\boldsymbol{x})=\varphi^*f(\boldsymbol{x})$, 如图 1.3.3 所示.

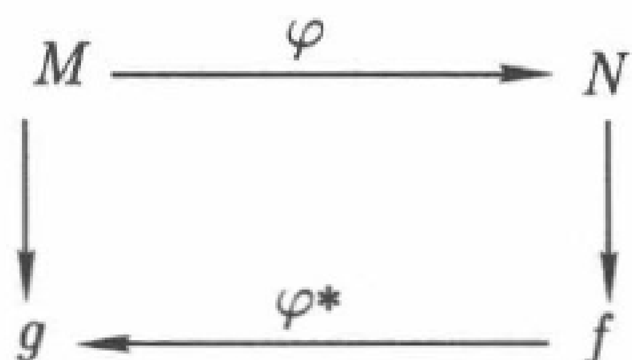


图 1.3.3 微分 0-形式的拉回映射

在选定流形的局部坐标后, 可以给出更具体的拉回映射表达式. 假设 $\omega_i(\boldsymbol{y})$ 是微分流形 N 上给定的微分 1-形式:

$$\omega_i(\boldsymbol{y})=a_i(\boldsymbol{y})\mathrm{d}y_i\in\Lambda^1(N),$$

可拉回得到微分流形 M 上的 1-形式

$$\varphi^*(a_i(\boldsymbol{y})\mathrm{d}y_i) = a_i(\varphi(\boldsymbol{x}))\sum_{k=1}^{m}\frac{\partial y_i}{\partial x_k}\mathrm{d}x_k \in \Lambda^1(M). \tag{1.3.5}$$

给定微分流形 N 上的 k-形式

$$\omega^k(\boldsymbol{y}) = \frac{1}{k!}a_{i_1\cdots i_k}(y)\mathrm{d}y_{i_1}\wedge\cdots\wedge\mathrm{d}y_{i_k},$$

它可拉回得到微分流形 M 上的 k-形式

$$\varphi^*(\omega^k(\boldsymbol{y})) = \frac{1}{k!}a_{i_1\cdots i_k}(\boldsymbol{y}(\boldsymbol{x})) \cdot \sum_{1\leqslant j_1<\cdots<j_k\leqslant m}\frac{\partial(y_{i_1},\cdots,y_{i_k})}{\partial(x_{j_1},\cdots,x_{j_k})}\mathrm{d}x_{j_1}\wedge\cdots\wedge\mathrm{d}x_{j_k}. \tag{1.3.6}$$

拉回映射具有下列性质:

(1) **线性性**:

$$\varphi^*(\alpha\omega_1+\beta\omega_2) = \alpha\varphi^*\omega_1+\beta\varphi^*\omega_2, \quad \omega_1,\omega_2\in\Lambda^k(N),\ \alpha,\beta\in\mathbf{R};$$

(2) **外积交换性**:

$$\varphi^*(\omega_1\wedge\omega_2) = \varphi^*\omega_1\wedge\varphi^*\omega_2;$$

(3) **连续性**:

$$\varphi: M\to N, \quad \psi: N\to L,$$
$$(\psi\circ\varphi)^* = \varphi^*\circ\psi^*.$$

定义 1.3.5　假设 $\varphi: M\to N$ 是一个 C^r 微分同胚, $X\in\chi(M)$. 由 φ 定义的 X 的**推前映射**为

$$\varphi_*X = T\varphi\circ X\circ\varphi^{-1}\in\chi(N), \tag{1.3.7}$$

其中 $T\varphi$ 是由 φ 诱导出的切映射.

定义 1.3.6　如果向量场 $X\in\chi(M)$ 和 $Y\in\chi(N)$ 满足 $T\varphi\circ X = Y\circ\varphi$, 则称 X 和 Y 是 φ-**相关的**, 记为 $X\overset{\varphi}{\sim}Y$ (图 1.3.4).

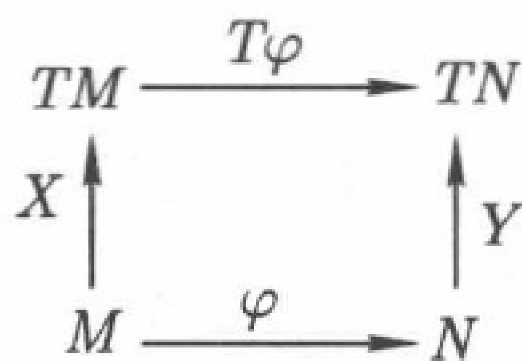


图 1.3.4 向量场 X 和 Y φ-相关

注 1.3.1 微分流形之间的映射 φ 会诱导出微分流形上张量场之间的映射. 对切向量场, 产生推前切映射 φ_*; 对微分形式, 导致方向相反的拉回映射 φ^*. 它们的对偶性表现为

$$(\varphi^*\omega, X)_{\boldsymbol{x}} = (\omega, \varphi_* X)_{\boldsymbol{y}=\varphi(\boldsymbol{x})}.$$

上式两端是在不同微分流形上的函数:

$$(\varphi^*\omega, X) \in \mathcal{F}(M), \quad (\omega, \varphi_* X) \in \mathcal{F}(N).$$

定义 1.3.7 微分形式 ω 称为**闭形式**, 如果 $\mathrm{d}\omega = 0$.

定义 1.3.8 微分 k-形式 ω 称为**恰当形式**, 如果存在 $(k-1)$-形式 η, 使得 $\omega = \mathrm{d}\eta$.

定理 1.3.1 如果 ω 是恰当的, 则 ω 是闭的; 反之不然.

定理 1.3.2 (Poincaré 引理) 在星形区域上, 闭微分形式即是恰当微分形式.

1.3.4 李导数和缩并

设 $\varphi: \mathbf{R}^m \to \mathbf{R}^n$ 是一个微分同胚, 且 $Y \in \mathcal{X}(\mathbf{R}^n)$ 是 $\mathbf{R}^n$ 上的一个光滑向量场.

定义 1.3.9 向量场 Y 的**拉回** φ^*Y 是 $\mathbf{R}^m$ 上的光滑向量场, 满足下面的表达式:

$$(\varphi^*Y)(\boldsymbol{x}) = (\mathrm{D}\varphi^{-1})(\boldsymbol{y})Y(\boldsymbol{y}) = \varphi_*^{-1}(\boldsymbol{y})Y(\boldsymbol{y}), \quad \boldsymbol{y} = \varphi(\boldsymbol{x}),$$

其中 $\mathrm{D} = \dfrac{\partial}{\partial \boldsymbol{y}}$.

定义 1.3.10 设 X 和 Y 是 $\mathbf{R}^n$ 上的两个光滑向量场, Y 对 X 的**李导数**定义为

$$\mathrm{L}_X Y = \frac{\mathrm{d}}{\mathrm{d}t}(\phi_X^{t*} Y)\Big|_{t=0},$$

其中 ϕ_X^t 是 X 的流.

定理 1.3.3 设 X 和 Y 是 $\mathbf{R}^n$ 上的两个光滑向量场, 则有

(1) $(\phi_X^{t*}Y)f = \phi_X^{t^*}(Y\phi_X^{-t^*}f), \forall f \in C^\infty(\mathbf{R}^n)$;

(2) $\dfrac{\mathrm{d}}{\mathrm{d}t}(\phi_X^{t^*}f) = \phi_X^{t^*}(Xf), \forall f \in C^\infty(\mathbf{R}^n)$;

(3) $\mathrm{L}_X Y = [X, Y]$, 其中 $[\cdot,\cdot]$ 为李括号.

证明 (1) 由定义 1.2.10 得到

$$
\begin{aligned}
(\phi_X^{t^*}Y)f(\boldsymbol{x}) &= \langle \mathrm{d}f(\boldsymbol{x}), (\phi_X^{t^*}Y)(\boldsymbol{x})\rangle = \langle \mathrm{d}f(\boldsymbol{x}), \mathrm{D}\phi_X^{-t}(\boldsymbol{y})Y(\boldsymbol{y})\rangle \quad (\boldsymbol{y} = \phi_X^t(\boldsymbol{x}))\\
&= \langle (\mathrm{D}\phi_X^{-t}(\boldsymbol{y}))^{\mathrm{T}}\mathrm{d}f(\boldsymbol{x}), Y(\boldsymbol{y})\rangle\\
&= \langle \mathrm{d}(f\circ\phi_X^{-t}(\boldsymbol{y}))^{\mathrm{T}}, Y(\boldsymbol{y})\rangle\\
&= \langle \mathrm{d}\widetilde{f}(\boldsymbol{y}), Y(\boldsymbol{y})\rangle \quad (\widetilde{f} = f\circ\phi_X^{-t})\\
&= (Y\widetilde{f})(\boldsymbol{y}) = (Y\widetilde{f})(\phi_X^t(\boldsymbol{x}))\\
&= \phi_X^{t^*}(Y\widetilde{f})(\boldsymbol{x}) = \phi_X^{t^*}(Yf(\phi_X^{-t}(\boldsymbol{y})))(\boldsymbol{x})\\
&= \phi_X^{t^*}(Y(\phi_X^{-t^*}f))(\boldsymbol{x}).
\end{aligned}
$$

(2)
$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}(\phi_X^{t^*}f) &= \frac{\mathrm{d}}{\mathrm{d}t}f(\phi_X^t(\boldsymbol{x})) = \sum_{i=1}^n \frac{\partial f}{\partial x^i}X_i(\phi_X^t(\boldsymbol{x}))\\
&= (Xf)(\phi_X^t(\boldsymbol{x})) = \phi_X^{t^*}(Xf),
\end{aligned}
$$

其中 X_i 是 X 的局部坐标.

(3) 因为对 $\forall f \in C^\infty(\mathbf{R}^n)$, 有

$$
\begin{aligned}
(\mathrm{L}_X Y)f &= \frac{\mathrm{d}}{\mathrm{d}t}(\phi_X^{t^*}Y)f\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}(\phi_X^{t^*}(Y(\phi_X^{-t^*}f)))\Big|_{t=0}\\
&= ((\phi_X^{t^*}X)Y(\phi_X^{-t^*}f) - \phi_X^{t^*}Y\phi_X^{-t^*}Xf)\big|_{t=0}\\
&= XYf - YXf = [X, Y]f,
\end{aligned}
$$

所以

$$
\mathrm{L}_X Y = [X, Y]. \qquad \square
$$

与 (3) 等价的就是关于李括号 $[\cdot,\cdot]$ 的 Jacobi 恒等式. 由 (3) 知道, 对任意光滑向量场 $X, Y, Z \in \mathcal{X}(\mathbf{R}^n)$, 有

$$
\mathrm{L}_X[Y, Z] = [\mathrm{L}_X Y, Z] + [Y, \mathrm{L}_X Z].
$$

上式显然等价于 Jacobi 恒等式.

定义 1.3.11 对 $\forall\omega\in\Lambda^k(\mathbf{R}^n)$, ω 对向量场 $X\in\mathcal{X}(\mathbf{R}^n)$ 的**李导数** $\mathrm{L}_X\omega$ 定义如下:

$$\mathrm{L}_X\omega=\frac{\mathrm{d}}{\mathrm{d}t}\phi_X^{t^*}\omega\Big|_{t=0}=\lim_{t\to 0}\frac{1}{t}(\phi_X^{t^*}\omega-\omega).$$

定理 1.3.4 对向量场 $X\in\mathcal{X}(\mathbf{R}^n)$ 的李导数, 有如下性质:

(1) $\mathrm{L}_Xf=Xf=\sum_{i=1}^{n}X_i\dfrac{\partial f}{\partial x^i}$, $f\in C^\infty(\mathbf{R}^n)$, 即函数 f 对向量场 X 的李导数是 f 沿方向 X 的方向导数;

(2) L_X 是一个**斜导数**, 即对 $\forall\omega_1,\omega_2\in\Lambda^k(\mathbf{R}^n)$, 有

$$\mathrm{L}_X(\alpha\omega_1+\beta\omega_2)=\alpha\mathrm{L}_X\omega_1+\beta\mathrm{L}_X\omega_2,\quad \alpha,\beta\in\mathbf{R},$$
$$\mathrm{L}_X(\omega_1\wedge\omega_2)=\mathrm{L}_X\omega_1\wedge\omega_2+\omega_1\wedge\mathrm{L}_X\omega_2;$$

(3) $\mathrm{L}_X\mathrm{d}=\mathrm{d}\mathrm{L}_X$.

证明 (1) 因为 $\dfrac{\mathrm{d}}{\mathrm{d}t}\phi_X^{t^*}f=\phi_X^{t^*}(Xf)$, 所以

$$\mathrm{L}_Xf=\frac{\mathrm{d}}{\mathrm{d}t}(\phi_X^{t^*}f)\Big|_{t=0}=(\phi_X^{t^*}Xf)\big|_{t=0}=Xf.$$

(2) 显然 L_X 是双线性的, 又有

$$\begin{aligned}\mathrm{L}_X(\omega_1\wedge\omega_2)&=\frac{\mathrm{d}}{\mathrm{d}t}\phi_X^{t^*}(\omega_1\wedge\omega_2)\Big|_{t=0}\\&=\frac{\mathrm{d}}{\mathrm{d}t}(\phi_X^{t^*}\omega_1\wedge\phi_X^{t^*}\omega_2)\Big|_{t=0}\\&=\left(\frac{\mathrm{d}}{\mathrm{d}t}\phi_X^{t^*}\omega_1\Big|_{t=0}\right)\wedge\omega_2+\omega_1\wedge\frac{\mathrm{d}}{\mathrm{d}t}\phi_X^{t^*}\omega_2\Big|_{t=0}\\&=\mathrm{L}_X\omega_1\wedge\omega_2+\omega_1\wedge\mathrm{L}_X\omega_2.\end{aligned}$$

(3) 由于对 $\forall\omega\in\Lambda^k(\mathbf{R}^n)$, 有

$$\begin{aligned}\mathrm{L}_X\mathrm{d}\omega&=\frac{\mathrm{d}}{\mathrm{d}t}\phi_X^{t^*}\mathrm{d}\omega\Big|_{t=0}=\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{d}\phi_X^{t^*}\omega\Big|_{t=0}\\&=\mathrm{d}\frac{\mathrm{d}}{\mathrm{d}t}\phi_X^{t^*}\omega\Big|_{t=0}=\mathrm{d}\mathrm{L}_X\omega,\end{aligned}$$

所以 $\mathrm{L}_X\mathrm{d}=\mathrm{d}\mathrm{L}_X$. □

定义 1.3.12 设 $X \in \mathcal{X}(\mathbf{R}^n)$, 且 $\omega \in \Lambda^k(\mathbf{R}^n)$, ω 和 X 的**缩并** (或称为 ω 与向量场 X 的**内积**) $i_X\omega$ 定义如下:

$$i_X\omega(\xi_1, \cdots, \xi_{k-1}) = \omega(X, \xi_1, \cdots, \xi_{k-1}), \quad \xi_i \in T_x\mathbf{R}^n,\ i = 1, \cdots, k-1.$$

显然上式定义了一个从 k-形式到 $(k-1)$-形式的映射 i_X, 即

$$i_X : \Lambda^k(\mathbf{R}^n) \to \Lambda^{k-1}(\mathbf{R}^n).$$

特别地, $i_X f = 0, \forall f \in C^\infty(\mathbf{R}^n) = \Lambda^0(\mathbf{R}^n)$.

例 1.3.3 设 $(x, y, z) \in \mathbf{R}^3$, 则有

$$i_{\partial_x}(\mathrm{d}x \wedge \mathrm{d}y) = \mathrm{d}y, \quad i_{\partial_x}(\mathrm{d}z \wedge \mathrm{d}x) = -\mathrm{d}z, \quad i_{\partial_x}(\mathrm{d}y \wedge \mathrm{d}z) = 0,$$

其中 $\partial_x = \dfrac{\partial}{\partial x}$.

注 1.3.2 如果 θ 为 $\mathbf{R}^n$ 上的任意微分形式, $\omega \in \Lambda^k(\mathbf{R}^n)$, 则有

$$i_\nu(\omega \wedge \theta) = (i_\nu\omega) \wedge \theta + (-1)^k\omega \wedge (i_\nu\theta), \quad \nu \in \mathcal{X}(\mathbf{R}^n).$$

定理 1.3.5 设 $\omega_1 \in \Lambda^k(\mathbf{R}^n)$, $\omega_2 \in \Lambda^l(\mathbf{R}^n)$, 则

(1) i_X 具有下面的性质:

$$i_X(\alpha_1\omega_1 + \alpha_2\omega_2) = \alpha_1 i_X\omega_1 + \alpha_2 i_X\omega_2, \quad \alpha_1, \alpha_2 \in \mathbf{R},$$
$$i_X(\omega_1 \wedge \omega_2) = i_X\omega_1 \wedge \omega_2 + (-1)^k\omega_1 \wedge i_X\omega_2;$$

(2) $i_{fX+gY} = fi_X + gi_Y,\ \forall f, g \in C^\infty(\mathbf{R}^n),\ X, Y \in \mathcal{X}(\mathbf{R}^n)$;

(3) $i_X\mathrm{d}f = \mathrm{L}_X f,\ f \in C^\infty(\mathbf{R}^n)$;

(4) $\mathrm{L}_X = i_X\mathrm{d} + \mathrm{d}i_X$;

(5) $\mathrm{L}_{fX} = f\mathrm{L}_X + \mathrm{d}f \wedge i_X$.

定理根据李导数的性质和定义 1.3.2 即可得证.

定理 1.3.6 设 $\omega \in \Lambda^k(\mathbf{R}^n)$, X 和 $X_i\,(i = 0, \cdots, k)$ 是 $\mathbf{R}^n$ 上的向量场, 于是有

(1) $$(\mathrm{L}_X\omega)(X_1, \cdots, X_k) = \mathrm{L}_X(\omega(X_1, \cdots, X_k)) - \sum_{i=1}^{k}\omega(X_1, \cdots, \mathrm{L}_X X_i, \cdots, X_k).$$

(2) $\mathrm{d}\omega(X_0, X_1, \cdots, X_k)$

$$= \sum_{i=0}^{k}(-1)^i \mathrm{L}_{X_i}(\omega(X_0, \cdots, \widehat{X}_i, \cdots, X_k))$$
$$+ \sum_{i<j}(-1)^{i+j}\omega(\mathrm{L}_{X_i}X_j, X_0, \cdots, \widehat{X}_i, \cdots, \widehat{X}_j, \cdots, X_k),$$

其 $\widehat{X}_i$ 和 $\widehat{X}_j$ 分别表示不含 X_i 和 X_j 项.

证明 (1) 根据微分形式拉回映射的定义知道

$$\begin{aligned}(\phi_X^{t^*}\omega)(X_1, \cdots, X_k) &= \omega(\phi_X^t(\boldsymbol{x}))(\phi_{X_*}^t X_1, \cdots, \phi_{X_*}^t X_k) \\ &= \phi_X^{t^*}(\omega(\phi_X^{-t^*} X_1, \cdots, \phi_X^{-t^*} X_k))(\phi_X^t(\boldsymbol{x})) \\ &= \phi_X^{t^*}(\omega(\phi_X^{-t^*} X_1, \cdots, \phi_X^{-t^*} X_k)),\end{aligned}$$

再根据李导数定义 1.3.11 及其性质得到

$$\begin{aligned}(\mathrm{L}_X\omega)(X_1, \cdots, X_k) &= \frac{\mathrm{d}}{\mathrm{d}t}(\phi_X^{t^*}\omega)(X_1, \cdots, X_k)\Big|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}(\phi_X^{t^*}\omega(\phi_X^{-t^*} X_1, \cdots, \phi_X^{-t^*} X_k))\Big|_{t=0} \\ &= \mathrm{L}_X(\omega(X_1, \cdots, X_k)) - \sum_{i=1}^{k}\omega(X_1, \cdots, \mathrm{L}_X X_i, \cdots, X_k).\end{aligned}$$

(2) 用归纳法证明. 当 $k = 0$ 时, 显然成立, 因为 $\mathrm{d}f(X) = \mathrm{L}_X f, f \in C^\infty(\mathbf{R}^n) = \Lambda^0(\mathbf{R}^n)$. 设 (2) 对 $k-1$ 为真, 下面证明对 k 也真. 由 (1) 得

$$\begin{aligned}\mathrm{d}\omega(X_0, X_1, \cdots, X_k) &= (i_{X_0}\mathrm{d}\omega)(X_1, \cdots, X_k) \\ &= (\mathrm{L}_{X_0}\omega)(X_1, \cdots, X_k) - (\mathrm{d}i_{X_0}\omega)(X_1, \cdots, X_k) \\ &= \mathrm{L}_{X_0}(\omega(X_1, \cdots, X_k)) - \sum_{i=1}^{k}\omega(X_1, \cdots, \mathrm{L}_{X_0}X_i, \cdots, X_k) \\ &\quad -(\mathrm{d}i_{X_0}\omega)(X_1, \cdots, X_k).\end{aligned}$$

这里 $i_{X_0}\omega \in \Lambda^{k-1}(\mathbf{R}^n)$, 由归纳假设知

$$
\begin{aligned}
&(\mathrm{d}i_{X_0}\omega)(X_1,\cdots,X_k)\\
&\quad=\sum_{i=1}^{k}(-1)^{i-1}\mathrm{L}_{X_i}(i_{X_0}\omega(X_1,\cdots,\widehat{X}_i,\cdots,X_k))\\
&\qquad+\sum_{1\leqslant i<j\leqslant k}(-1)^{i+j}i_{X_0}\omega(\mathrm{L}_{X_i}X_j,X_1,\cdots,\widehat{X}_i,\cdots,\widehat{X}_j,\cdots,X_k)\\
&\quad=\sum_{i=1}^{k}(-1)^{i-1}L_{X_i}(\omega(X_0,X_1,\cdots,\widehat{X}_i,\cdots,X_k))\\
&\qquad+\sum_{1\leqslant i<j\leqslant k}(-1)^{i+j-1}\omega(\mathrm{L}_{X_i}X_j,X_0,\cdots,\widehat{X}_i,\cdots,\widehat{X}_j,\cdots,X_k).
\end{aligned}
$$

这样, 我们可得

$$
\begin{aligned}
&\mathrm{d}\omega(X_0,X_1,\cdots,X_k)\\
&\quad=\mathrm{L}_{X_0}(\omega(X_1,\cdots,X_k))+\sum_{j=1}^{k}(-1)^{j}\omega(X_1,\cdots,\mathrm{L}_{X_0}X_j,\cdots,X_k)\\
&\qquad+\sum_{i=1}^{k}(-1)^{i}\mathrm{L}_{X_i}(\omega(X_0,\cdots,\widehat{X}_i,\cdots,X_k))\\
&\qquad+\sum_{1\leqslant i<j<k}(-1)^{i+j}\omega(\mathrm{L}_{X_i}X_j,X_0,\cdots,\widehat{X}_i,\cdots,\widehat{X}_j,\cdots,X_k)\\
&\quad=\sum_{i=0}^{k}(-1)^{i}\mathrm{L}_{X_i}(\omega(X_0,\cdots,\widehat{X}_i,\cdots,X_k))\\
&\qquad+\sum_{i<j}(-1)^{i+j}\omega(\mathrm{L}_{X_i}X_j,X_0,\cdots,\widehat{X}_i,\cdots,\widehat{X}_j,\cdots,X_k).
\end{aligned}
$$

□

参 考 文 献

[San83] Santilli R M. *Foundations of Theoretical Mechanics I, II.* 2nd edition. New York: Springer-Verlag, 1983.

[Arn89] Arnold V I. *Mathematical Methods of Classical Mechanics.* 2nd edition. Berlin: Springer-Verlag, 1989.

[FQ09] Feng K, Qin M Z. *Symplectic Geometric Algorithms for Hamiltonian Systems.* Zhejiang: Springer-Verlag GmbH, 2009.

[FQ91a] Feng K, Qin M Z. Hamiltonian algorithms for Hamiltonian dynamical systems. *Prog Nat Sci*, 1991, 1(2): 105–116.

[Kol54] Kolmogorov A N. On conservation of conditionally periodic motions under small perturbations of the Hamiltonian. *Dokl Akad Nauk SSSR*, 1954, 98: 527–530.

[Arn63] Arnold V I. Proof of a theorem of A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian. *Russ Math Surv*, 1963, 18(5): 9–36.

[Mos62] Moser J. On invariant curves of area-preserving mappings of an annulus. *Nachr Akad Wiss Gottingen II Math Phys*, 1962: 1–20.

[Shang99] Shang Z. KAM theorem of symplectic algorithms for Hamiltonian systems. *Numer Math*, 1999, 83: 477–496.

第二章 Hamilton 系统的辛差分格式

本章主要介绍几类常见的 Hamilton 辛算法及其性质.

Hamilton 系统具有许多内在的特性, 比如它的相流保持相空间的面积和体积不变, 保持形式能量和动量不变等, 特别是相流保持相空间的辛结构不变, 这也是 Hamilton 系统的重要特征之一. 事实上, 辛结构不变本身就隐含了相空间的面积和体积的不变性. 因此, 我们在数值求解 Hamilton 方程时, 也希望能保持这一特征. 能够保持辛结构不变这一特性的差分格式就是辛格式, 即从上一步到下一步的迭代映射 (也称为步进映射) 是辛映射.

设 H 是 $2n$ 个变量 $p_1, \cdots, p_n, q_1, \cdots, q_n$ 的可微函数, 则 Hamilton 方程是

$$\dot{\boldsymbol{p}} = -H_{\boldsymbol{q}}, \quad \dot{\boldsymbol{q}} = H_{\boldsymbol{p}}, \tag{2.0.1}$$

其中 $\boldsymbol{p} = (p_1, \cdots, p_n)^{\mathrm{T}}, \boldsymbol{q} = (q_1, \cdots, q_n)^{\mathrm{T}}$. 为了表达方便, 我们记

$$\boldsymbol{z} = \begin{pmatrix} p_1 \\ \vdots \\ p_n \\ q_1 \\ \vdots \\ q_n \end{pmatrix} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \\ z_{n+1} \\ \vdots \\ z_{2n} \end{pmatrix}, \quad H_{\boldsymbol{z}} = \begin{pmatrix} \dfrac{\partial H}{\partial z_1} \\ \vdots \\ \dfrac{\partial H}{\partial z_{2n}} \end{pmatrix},$$

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I}_n \\ -\boldsymbol{I}_n & \boldsymbol{0} \end{pmatrix}, \tag{2.0.2}$$

其中 $\boldsymbol{I}_n$ 是 n 阶单位矩阵, $\boldsymbol{J}$ 的逆矩阵 $\boldsymbol{J}^{-1} = \boldsymbol{J}^{\mathrm{T}} = -\boldsymbol{J}$. 利用这些记号, 方程 (2.0.1) 可以表示为

$$\frac{\mathrm{d}\boldsymbol{z}}{\mathrm{d}t} = \boldsymbol{J}^{-1} H_{\boldsymbol{z}}. \tag{2.0.3}$$

函数 H 称为该系统的 **Hamilton 函数**.

用 g_H^t 表示方程 (2.0.1) 的相流 (指向量场 $\boldsymbol{J}^{-1}H_{\boldsymbol{z}}$ 的相流), 后面简记为 g^t. 由 Hamilton 力学的基本定理知道, 任何一个 Hamilton 系统的解是一个单参数辛群, 用 $Sp(2n)$ 来表示. 由此可见, Hamilton 力学是建立在辛几何基础上的. 因为局部地看每一个 $2n$ 维流形都同胚于 $\mathbf{R}^{2n}$ 的质点的邻域, 所以为了简单起见, 我们仅考虑经典的相空间 $\mathbf{R}^{2n}_{\boldsymbol{p}\times\boldsymbol{q}} = \mathbf{R}^n_{\boldsymbol{p}} \times \mathbf{R}^n_{\boldsymbol{q}}$, 其中 $\mathbf{R}^n_{\boldsymbol{p}}$ 为动量空间, $\mathbf{R}^n_{\boldsymbol{q}}$ 为构形空间.

Hamilton 系统相空间配备着一个标准辛形式或称标准辛结构, 也就是一个闭的微分 2-形式

$$\omega_{\boldsymbol{J}} = \sum_{i=1}^{n} \mathrm{d}z_i \wedge \mathrm{d}z_{n+i} = \sum_{i=1}^{n} \mathrm{d}p_i \wedge \mathrm{d}q_i, \tag{2.0.4}$$

即对于 $\mathbf{R}^{2n}$ 中的每一点 $\boldsymbol{z}$, 它是点 $\boldsymbol{z}$ 的切空间 $T_{\boldsymbol{z}}\mathbf{R}^{2n}$ 上的反对称双线性形式:

$$\omega_{\boldsymbol{J}}(\boldsymbol{\xi}, \boldsymbol{\eta}) = \boldsymbol{\xi}^{\mathrm{T}} \boldsymbol{J} \boldsymbol{\eta}, \quad \forall \boldsymbol{\xi}, \boldsymbol{\eta} \in T_{\boldsymbol{z}}\mathbf{R}^{2n},$$

其中 $\boldsymbol{J}$ 是标准反对称矩阵 (2.0.2).

令 $\boldsymbol{w} : \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ 是一个从点 $\boldsymbol{z} \in \mathbf{R}^{2n}$ 到点 $\boldsymbol{w}(\boldsymbol{z}) \in \mathbf{R}^{2n}$ 的可微映射, 则其 Jacobi 矩阵为

$$\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{z}} = \begin{pmatrix} \dfrac{\partial w_1}{\partial z_1} & \cdots & \dfrac{\partial w_1}{\partial z_{2n}} \\ \vdots & & \vdots \\ \dfrac{\partial w_{2n}}{\partial z_1} & \cdots & \dfrac{\partial w_{2n}}{\partial z_{2n}} \end{pmatrix}.$$

对于每一个点 $\boldsymbol{z} \in \mathbf{R}^{2n}$, 映射 $\boldsymbol{w}$ 诱导出一个从点 $\boldsymbol{z}$ 的切空间到点 $\boldsymbol{w}(\boldsymbol{z})$ 的切空间的线性映射 $\boldsymbol{w}_*(\boldsymbol{z})$, 即

$$\boldsymbol{\xi} = (\xi_1, \cdots, \xi_{2n})^{\mathrm{T}} \mapsto \boldsymbol{w}_*\boldsymbol{\xi} = \frac{\partial \boldsymbol{w}}{\partial \boldsymbol{z}}\boldsymbol{\xi}, \quad \boldsymbol{\xi} \in T_{\boldsymbol{z}}\mathbf{R}^{2n}.$$

对于每一个 $\mathbf{R}^{2n}$ 上的 2-形式 ω, 它同样诱导出一个在 $\mathbf{R}^{2n}$ 上的 2-形式 $\boldsymbol{w}^*\omega$:

$$\boldsymbol{w}^*\omega(\boldsymbol{\xi}, \boldsymbol{\eta})_{\boldsymbol{z}} = \omega\left(\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{z}}\boldsymbol{\xi}, \frac{\partial \boldsymbol{w}}{\partial \boldsymbol{z}}\boldsymbol{\eta}\right)_{\boldsymbol{w}(\boldsymbol{z})}, \quad \boldsymbol{\xi}, \boldsymbol{\eta} \in T_{\boldsymbol{z}}\mathbf{R}^{2n},$$

见定义 1.3.4.

若 $\omega(\boldsymbol{\xi},\boldsymbol{\eta})_{\boldsymbol{z}}=\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{z})\boldsymbol{\eta}$, $\boldsymbol{A}(\boldsymbol{z})^{\mathrm{T}}=-\boldsymbol{A}(\boldsymbol{z})$, 则有

$$\boldsymbol{w}^{*}\omega(\boldsymbol{\xi},\boldsymbol{\eta})_{\boldsymbol{z}}=\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{z})\boldsymbol{\eta},$$

其中

$$\boldsymbol{B}(\boldsymbol{z})=\left(\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{z}}\right)^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{w}(\boldsymbol{z}))\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{z}}.$$

定义 2.0.1 假设 $\boldsymbol{w}$ 是 $\mathbf{R}^{2n}$ 上的一个微分同胚, 如果 $\boldsymbol{w}$ 保持标准辛结构, 即 $\boldsymbol{w}^{*}\omega_{\boldsymbol{J}}=\omega_{\boldsymbol{J}}$, 也就是

$$\left(\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{z}}\right)^{\mathrm{T}}\boldsymbol{J}\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{z}}=\boldsymbol{J}, \tag{2.0.5}$$

则称 $\boldsymbol{w}$ 是一个**正则变换**或**标准辛变换**. 标准辛变换 $\boldsymbol{w}$ 的 Jacobi 矩阵是辛矩阵.

正则变换的几何意义见图 2.0.1.

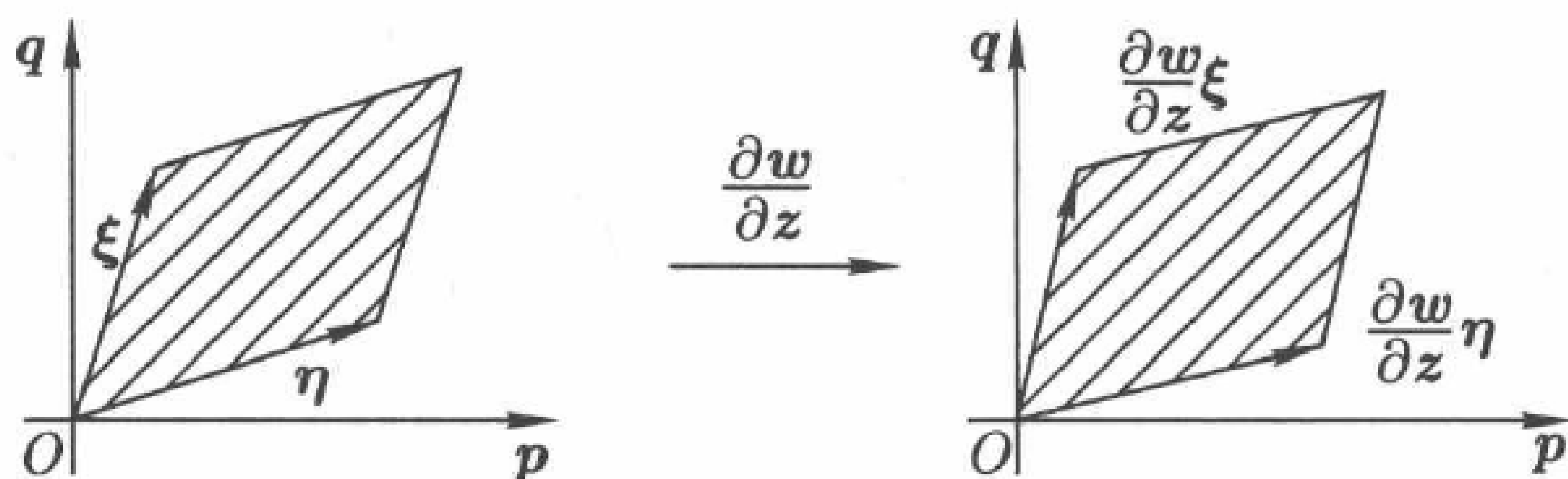


图 2.0.1 正则变换 $\boldsymbol{w}$ 的几何意义

按照常微分方程理论, 方程 (2.0.1) 至少在 $(\boldsymbol{z},t)$ 局部存在一个 $\mathbf{R}^{2n}$ 上的单参数群 g^t, 满足

$$g^0=\mathrm{id},\quad g^{t_1+t_2}=g^{t_1}\cdot g^{t_2}.$$

假设 $\boldsymbol{z}(t)$ 在 $t=0$ 处的初始值为 $\boldsymbol{z}_0$, 则方程 (2.0.1) 的解可写成

$$\boldsymbol{z}(t)=g^t\boldsymbol{z}_0.$$

从前面关于 Hamilton 方程基本性质的讨论知道, 对一切 t, g^t 是一个正则变换, 即

$$(g^t)^{*}\omega_{\boldsymbol{J}}=\omega_{\boldsymbol{J}}.$$

这就引出下列相空间面积守恒律:

$$\int_{g^t\sigma^2}\omega_{\boldsymbol{J}}=\int_{\sigma^2}\omega_{\boldsymbol{J}},\qquad 2\text{-链 }\sigma^2\subset\mathbf{R}^{2n},$$

$$\int_{g^t\sigma^4}\omega_{\boldsymbol{J}}\wedge\omega_{\boldsymbol{J}}=\int_{\sigma^4}\omega_{\boldsymbol{J}}\wedge\omega_{\boldsymbol{J}},\qquad 4\text{-链 }\sigma^4\subset\mathbf{R}^{2n},$$

$$\cdots\cdots$$

$$\int_{g^t\sigma^{2n}}\omega_{\boldsymbol{J}}\wedge\cdots\wedge\omega_{\boldsymbol{J}}=\int_{\sigma^{2n}}\omega_{\boldsymbol{J}}\wedge\cdots\wedge\omega_{\boldsymbol{J}},\quad 2n\text{-链 }\sigma^{2n}\subset\mathbf{R}^{2n}.$$

最后一个等式就是 Liouville 相体积守恒律.

另一些类型的守恒量是能量和所有首次积分. 一个光滑函数 $\varphi(\boldsymbol{z})$ 称为系统 (2.0.1) 的**首次积分**, 当且仅当对一切 $(\boldsymbol{z},t)$ 成立 $\varphi(g^t\boldsymbol{z})=\varphi(\boldsymbol{z})$ 等价于 $\{\varphi,H\}=0$, 这里 $\{\cdot,\cdot\}$ 为 Poisson 括号. 所以 Hamilton 函数 H 是对应 Hamilton 系统的首次积分. 上面的标准辛结构 (2.0.4) 能够推广到更一般的辛结构——一个非奇异的闭微分 2-形式

$$\omega_{\boldsymbol{K}}=\sum_{i<j}K_{ij}(\boldsymbol{z})\mathrm{d}z_i\wedge\mathrm{d}z_j,\tag{2.0.6}$$

即

$$\omega_{\boldsymbol{K}}(\boldsymbol{\xi},\boldsymbol{\eta})_{\boldsymbol{z}}=\frac{1}{2}\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{K}(\boldsymbol{z})\boldsymbol{\eta},\quad \boldsymbol{K}(\boldsymbol{z})^{\mathrm{T}}=-\boldsymbol{K}(\boldsymbol{z}),\quad \det\boldsymbol{K}(\boldsymbol{z})\neq 0,$$

其中 $\boldsymbol{K}(\boldsymbol{z})=(K_{ij}(\boldsymbol{z}))$.

定义 2.0.2 ($\boldsymbol{K}(z)$-辛变换) 一个微分同胚 $\boldsymbol{w}:\mathbf{R}^{2n}\to\mathbf{R}^{2n}$ 称为 **$\boldsymbol{K}(z)$-正则的**或 **$\boldsymbol{K}(z)$-辛的**, 如果 $\boldsymbol{w}^*\omega_{\boldsymbol{K}}=\omega_{\boldsymbol{K}}$, 即

$$\left(\frac{\partial\boldsymbol{w}}{\partial\boldsymbol{z}}\right)^{\mathrm{T}}\boldsymbol{K}(\boldsymbol{w}(\boldsymbol{z}))\frac{\partial\boldsymbol{w}}{\partial\boldsymbol{z}}=\boldsymbol{K}(\boldsymbol{z}).\tag{2.0.7}$$

Darboux 定理告诉我们所有的辛结构是等价的: 通过合适的坐标变换 $\boldsymbol{z}\mapsto\boldsymbol{w}(\boldsymbol{z})$, 一般辛结构 (2.0.6) 可以变为标准辛结构 (2.0.4), 即

$$\sum_{i<j}K_{ij}(\boldsymbol{z})\mathrm{d}z_i\wedge\mathrm{d}z_j=\sum_{i<j}\mathrm{d}w_i\wedge\mathrm{d}w_j.$$

§2.1　辛矩阵的一些性质

定义 2.1.1　一个 $2n$ 阶矩阵 $\boldsymbol{S}$ 是**辛矩阵** (简称为**辛阵**), 如果

$$\boldsymbol{S}^{\mathrm{T}}\boldsymbol{J}\boldsymbol{S}=\boldsymbol{J}. \tag{2.1.1}$$

所有辛阵组成的群称为辛群, 用 $Sp(2n)$ 来表示.

定义 2.1.2　一个 $2n$ 阶矩阵 $\boldsymbol{B}$ 称为**无穷小辛阵**, 如果

$$\boldsymbol{J}\boldsymbol{B}+\boldsymbol{B}^{\mathrm{T}}\boldsymbol{J}=\boldsymbol{0}. \tag{2.1.2}$$

所有无穷小辛阵配备运算 $[\boldsymbol{A},\boldsymbol{B}]=\boldsymbol{AB}-\boldsymbol{BA}$ 构成一个李代数, 用 $sp(2n)$ 来表示, 它是辛群 $Sp(2n)$ 的李代数.

下面不加证明地给出一些命题, 它们将直接或间接用于 Hamilton 系统辛算法的构造.

命题 2.1.1　若 $\boldsymbol{S}\in Sp(2n)$, 则 $|\boldsymbol{S}|=1$.

命题 2.1.2　若 $\boldsymbol{S}\in Sp(2n)$, 则 $\boldsymbol{S}^{-1}=-\boldsymbol{J}\boldsymbol{S}^{\mathrm{T}}\boldsymbol{J}=\boldsymbol{J}^{-1}\boldsymbol{S}^{\mathrm{T}}\boldsymbol{J}$.

命题 2.1.3　若 $\boldsymbol{S}\in Sp(2n)$, 则 $\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}^{\mathrm{T}}=\boldsymbol{J}$.

命题 2.1.4　设 $\boldsymbol{S}=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{pmatrix}$, 其中 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{D}$ 是 n 阶矩阵, 则 $\boldsymbol{S}$ 为辛阵的必要充分条件是

$$\boldsymbol{A}\boldsymbol{B}^{\mathrm{T}}-\boldsymbol{B}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{0},\quad \boldsymbol{C}\boldsymbol{D}^{\mathrm{T}}-\boldsymbol{D}\boldsymbol{C}^{\mathrm{T}}=\boldsymbol{0},\quad \boldsymbol{A}\boldsymbol{D}^{\mathrm{T}}-\boldsymbol{B}\boldsymbol{C}^{\mathrm{T}}=\boldsymbol{I},$$
$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{C}-\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{0},\quad \boldsymbol{B}^{\mathrm{T}}\boldsymbol{D}-\boldsymbol{D}^{\mathrm{T}}\boldsymbol{B}=\boldsymbol{0},\quad \boldsymbol{A}^{\mathrm{T}}\boldsymbol{D}-\boldsymbol{C}^{\mathrm{T}}\boldsymbol{B}=\boldsymbol{I}.$$

命题 2.1.5　矩阵

$$\begin{pmatrix}\boldsymbol{I} & \boldsymbol{B}\\ \boldsymbol{0} & \boldsymbol{I}\end{pmatrix},\quad \begin{pmatrix}\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{D} & \boldsymbol{I}\end{pmatrix}$$

是辛矩阵分别当且仅当 $\boldsymbol{B}^{\mathrm{T}}=\boldsymbol{B},\boldsymbol{D}^{\mathrm{T}}=\boldsymbol{D}$.

命题 2.1.6　矩阵 $\begin{pmatrix}\boldsymbol{A} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{D}\end{pmatrix}\in Sp(2n)$ 当且仅当 $\boldsymbol{A}=\boldsymbol{D}^{-\mathrm{T}}$.

命题 2.1.7　矩阵 $\boldsymbol{S}=\boldsymbol{M}^{-1}\boldsymbol{N}\in Sp(2n)$ 当且仅当

$$\boldsymbol{M}\boldsymbol{J}\boldsymbol{M}^{\mathrm{T}}=\boldsymbol{N}\boldsymbol{J}\boldsymbol{N}^{\mathrm{T}}.$$

命题 2.1.8 矩阵 $\begin{pmatrix} \boldsymbol{Q} & \boldsymbol{I}-\boldsymbol{Q} \\ -(\boldsymbol{I}-\boldsymbol{Q}) & \boldsymbol{Q} \end{pmatrix} \in Sp(2n)$ 当且仅当

$$\boldsymbol{Q}^2 = \boldsymbol{Q}, \quad \boldsymbol{Q}^{\mathrm{T}} = \boldsymbol{Q}.$$

命题 2.1.9 若 $\boldsymbol{B} \in sp(2n)$, 则 $\exp(\boldsymbol{B}) \in Sp(2n)$.

命题 2.1.10 若 $\boldsymbol{B} \in sp(2n)$, 且 $|\boldsymbol{I}+\boldsymbol{B}| \neq 0$, 则 $\boldsymbol{F} = (\boldsymbol{I}+\boldsymbol{B})^{-1}(\boldsymbol{I}-\boldsymbol{B}) \in Sp(2n)$, 称之为 $\boldsymbol{B}$ 的 **Cayley 变换**.

命题 2.1.11 若 $\boldsymbol{B} \in sp(2n)$, 则

$$(\boldsymbol{B}^{2m})^{\mathrm{T}}\boldsymbol{J} = \boldsymbol{J}(\boldsymbol{B}^{2m}), \quad m \in \mathbf{Z}^+.$$

命题 2.1.12 若 $\boldsymbol{B} \in sp(2n)$, 则

$$(\boldsymbol{B}^{2m+1})^{\mathrm{T}}\boldsymbol{J} = -\boldsymbol{J}(\boldsymbol{B}^{2m+1}), \quad m \in \mathbf{Z}^+.$$

命题 2.1.13 若 $f(x)$ 是偶次多项式, 且 $\boldsymbol{B} \in sp(2n)$, 则

$$f(\boldsymbol{B}^{\mathrm{T}})\boldsymbol{J} = \boldsymbol{J}f(\boldsymbol{B}).$$

命题 2.1.14 若 $h(x)$ 是奇次多项式, 且 $\boldsymbol{B} \in sp(2n)$, 则 $h(\boldsymbol{B}) \in sp(2n)$, 即

$$h(\boldsymbol{B}^{\mathrm{T}})\boldsymbol{J} + \boldsymbol{J}h(\boldsymbol{B}) = \boldsymbol{0}.$$

§2.2 线性 Hamilton 系统的辛格式

Hamilton 系统 (2.0.1) 称为**线性**的, 如果 Hamilton 函数 $H(\boldsymbol{z})$ 是 $\boldsymbol{z}$ 的二次型:

$$H(\boldsymbol{z}) = \frac{1}{2}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{z}, \quad \boldsymbol{C}^{\mathrm{T}} = \boldsymbol{C},$$

且 $\boldsymbol{J}$ 是一个标准反对称阵:

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I}_n \\ -\boldsymbol{I}_n & \boldsymbol{0} \end{pmatrix}, \quad \boldsymbol{J}^{\mathrm{T}} = -\boldsymbol{J} = \boldsymbol{J}^{-1}, \quad |\boldsymbol{J}| = 1.$$

于是正则方程 (2.0.1), (2.0.3) 变成

$$\frac{\mathrm{d}\boldsymbol{z}}{\mathrm{d}t} = \boldsymbol{B}\boldsymbol{z}, \quad \boldsymbol{B} = \boldsymbol{J}^{-1}\boldsymbol{C}, \quad \boldsymbol{C}^{\mathrm{T}} = \boldsymbol{C}, \tag{2.2.1}$$

其中 $\boldsymbol{B}$ 是无穷小辛阵. 方程 (2.2.1) 的解为

$$\boldsymbol{z}(t) = g^t \boldsymbol{z}(0), \quad g^t = \exp(t\boldsymbol{B}), \tag{2.2.2}$$

即 g^t 是无穷小辛阵 $t\boldsymbol{B}$ 的指数变换, 由命题 2.1.9 知它是辛阵.

现在考虑二次型 $F(\boldsymbol{z}) = \frac{1}{2}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{z}$. 两个二次型 H 和 F 的 Poisson 括号同样是二次型:

$$\{H, F\} = \frac{1}{2}\boldsymbol{z}^{\mathrm{T}}(\boldsymbol{AJC} - \boldsymbol{CJA})\boldsymbol{z}.$$

定理 2.2.1　二次型 F 是线性 Hamilton 系统 (2.2.1) 的不变积分的条件可表达为下述四个等价条件之一:

$$F((\exp(t\boldsymbol{J}^{-1}\boldsymbol{C}))\boldsymbol{z}) = F(\boldsymbol{z}), \tag{2.2.3}$$

$$\{H, F\} = 0, \tag{2.2.4}$$

$$(\exp(t\boldsymbol{J}^{-1}\boldsymbol{C}))^{\mathrm{T}}\boldsymbol{A}(\exp(t\boldsymbol{J}^{-1}\boldsymbol{C})) = \boldsymbol{A}, \tag{2.2.5}$$

$$\boldsymbol{AJC} = \boldsymbol{CJA}. \tag{2.2.6}$$

在文献 [Fen85] 中对方程 (2.2.1) 提出了一些辛格式. 第一个辛格式为 Euler 中点格式:

$$\frac{\boldsymbol{z}^{k+1} - \boldsymbol{z}^k}{\tau} = \boldsymbol{B}\frac{\boldsymbol{z}^{k+1} + \boldsymbol{z}^k}{2}, \tag{2.2.7}$$

其中 τ 为时间步长, 变换 $\boldsymbol{z}^k \mapsto \boldsymbol{z}^{k+1}$ 由下列关系式给出:

$$\boldsymbol{z}^{k+1} = \boldsymbol{F}_\tau \boldsymbol{z}^k, \quad \boldsymbol{F}_\tau = \phi\left(-\frac{\tau}{2}\boldsymbol{B}\right), \quad \phi(\lambda) = \frac{1-\lambda}{1+\lambda}. \tag{2.2.8}$$

显然, $\boldsymbol{F}_\tau$ 是无穷小辛阵 $-\frac{\tau}{2}\boldsymbol{B}$ 的 Cayley 变换, 根据命题 2.1.10 知它是辛阵.

第二个格式为可分 Hamilton 系统交叉显式辛格式. 称 Hamilton 系统为**可分的**, 若 $H(\boldsymbol{p}, \boldsymbol{q}) = U(\boldsymbol{p}) + V(\boldsymbol{q})$, 它的线性形式如下:

$$\begin{aligned} H(\boldsymbol{p}, \boldsymbol{q}) &= \frac{1}{2}(\boldsymbol{p}^{\mathrm{T}}, \boldsymbol{q}^{\mathrm{T}})\boldsymbol{S}\begin{pmatrix}\boldsymbol{p}\\ \boldsymbol{q}\end{pmatrix} = \frac{1}{2}\boldsymbol{p}^{\mathrm{T}}\boldsymbol{U}\boldsymbol{p} + \frac{1}{2}\boldsymbol{q}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{q} \\ &= U(\boldsymbol{p}) + V(\boldsymbol{q}), \end{aligned} \tag{2.2.9}$$

其中

$$\boldsymbol{S}=\begin{pmatrix}\boldsymbol{U} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{V}\end{pmatrix},$$

$\boldsymbol{U}^{\mathrm{T}}=\boldsymbol{U}$ 正定, 且 $\boldsymbol{V}^{\mathrm{T}}=\boldsymbol{V}$. 这时正则方程 (2.0.1) 变为

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}=-V_{\boldsymbol{q}},\quad \frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}t}=U_{\boldsymbol{p}}. \tag{2.2.10}$$

交叉显式辛格式为

$$\begin{aligned}&\frac{1}{\tau}(\boldsymbol{p}^{k+1}-\boldsymbol{p}^{k})=-V_{\boldsymbol{q}}^{k+\frac{1}{2}},\\ &\frac{1}{\tau}\left(\boldsymbol{q}^{k+\frac{1}{2}+1}-\boldsymbol{q}^{k+\frac{1}{2}}\right)=U_{\boldsymbol{p}}^{k+\frac{1}{2}},\end{aligned} \tag{2.2.11}$$

其中 $\boldsymbol{p}$ 在整时刻 $t=k\tau$ 上计算, 而 $\boldsymbol{q}$ 在半时刻 $t=\left(k+\dfrac{1}{2}\right)\tau$ 上计算. 变换

$$\boldsymbol{z}^{k}=\begin{pmatrix}\boldsymbol{p}^{k}\\ \boldsymbol{q}^{k+\frac{1}{2}}\end{pmatrix}\mapsto\begin{pmatrix}\boldsymbol{p}^{k+1}\\ \boldsymbol{q}^{k+\frac{1}{2}+1}\end{pmatrix}=\boldsymbol{z}^{k+1}$$

由下式给出:

$$\boldsymbol{z}^{k+1}=\boldsymbol{F}_{\tau}\boldsymbol{z}^{k},$$

其中

$$\boldsymbol{F}_{\tau}=\begin{pmatrix}\boldsymbol{I} & \boldsymbol{0}\\ -\boldsymbol{U} & \boldsymbol{I}\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{I} & -\boldsymbol{V}\\ \boldsymbol{0} & \boldsymbol{I}\end{pmatrix}. \tag{2.2.12}$$

由命题 2.1.5 易证上式右端乘积也是辛阵.

§2.3 基于 Padé 逼近的辛格式

如果轨道 $\boldsymbol{z}(t)=g^{t}\boldsymbol{z}_0$ 是满足初始条件 $\boldsymbol{z}(0)=\boldsymbol{z}_0$ 的线性 Hamilton 方程 (2.0.1) 的解, 相流 g^{t} 的 Jacobi 矩阵就是它本身 $\exp(t\boldsymbol{B})$. 那么, 逼近 $\exp(t\boldsymbol{B})$ 的最简单的办法就是利用 Padé 逼近. 我们首先考虑对 $\exp(x)$ 的有理逼近:

$$\exp(x)\sim\frac{n_{lm}(x)}{d_{lm}(x)}=g_{lm}(x), \tag{2.3.1}$$

其中

$$n_{lm}(x)=\sum_{k=0}^{m}\frac{(l+m-k)!m!}{(l+m)!k!(m-k)!}x^k, \tag{2.3.2}$$

$$d_{lm}(x)=\sum_{k=0}^{l}\frac{(l+m-k)!l!}{(l+m)!k!(l-k)!}(-x)^k. \tag{2.3.3}$$

对每一对非负整数 (l,m), 利用 Taylor 级数展开, 得到

$$\exp(x)-\frac{n_{lm}(x)}{d_{lm}(x)}=o(|x|^{m+l+1}),\quad |x|\to 0. \tag{2.3.4}$$

称 g_{lm} 为 $\exp(x)$ 的 $l+m$ 阶 **Padé 逼近**. 当 $l=m$ 时, 称 g_{ll} 为 **Padé 对角逼近**.

定理 2.3.1　设 $\boldsymbol{B}$ 是无穷小辛阵. 对于充分小的 $|t|$, $g_{lm}(t\boldsymbol{B})$ 是辛阵当且仅当 $l=m$, 即 $g_{ll}(x)$ 是 $\exp(x)$ 的 Padé 对角逼近.

证明　充分性　设 $n_{ll}(x)=f(x)+h(x)$, $d_{ll}=f(x)-h(x)$. 令 $f(x)$ 是偶次多项式, $h(x)$ 是奇次多项式. 为了证明 $g_{ll}(t\boldsymbol{B})\in Sp(2n)$, 我们仅需证明下面的等式成立:

$$(f(t\boldsymbol{B})+h(t\boldsymbol{B}))\boldsymbol{J}(f(t\boldsymbol{B})+h(t\boldsymbol{B}))^{\mathrm{T}}=(f(t\boldsymbol{B})-h(t\boldsymbol{B}))\boldsymbol{J}(f(t\boldsymbol{B})-h(t\boldsymbol{B}))^{\mathrm{T}}. \tag{2.3.5}$$

由命题 2.1.13 和命题 2.1.14 知, (2.3.5) 式的左边为

$$(f(t\boldsymbol{B})+h(t\boldsymbol{B}))\boldsymbol{J}(f(t\boldsymbol{B}^{\mathrm{T}})+h(t\boldsymbol{B}^{\mathrm{T}}))=(f(t\boldsymbol{B})+h(t\boldsymbol{B}))(f(t\boldsymbol{B})-h(t\boldsymbol{B}))\boldsymbol{J}. \tag{2.3.6}$$

同样, (2.3.5) 式的右边为

$$(f(t\boldsymbol{B})-h(t\boldsymbol{B}))\boldsymbol{J}(f(t\boldsymbol{B}^{\mathrm{T}})-h(t\boldsymbol{B}^{\mathrm{T}}))=(f(t\boldsymbol{B})-h(t\boldsymbol{B}))(f(t\boldsymbol{B})+h(t\boldsymbol{B}))\boldsymbol{J}. \tag{2.3.7}$$

比较 (2.3.6) 式和 (2.3.7) 式, 它们的右边显然相等, 因此左边也相等, 所以 (2.3.5) 式成立.

必要性　不失一般性, 假设 $l>m$. 因多项式 $d_{lm}(t\boldsymbol{B})\boldsymbol{J}d_{lm}(t\boldsymbol{B})$ 的阶高于多项式 $n_{lm}(t\boldsymbol{B})\boldsymbol{J}n_{lm}(t\boldsymbol{B})$ 的阶, 故 (2.3.5) 式不能成立.　□

下面在表 2.3.1 中给出 $\exp(-x)$ 的 Padé 逼近中前面的几项. 从 $\exp(-x)$ 的 Padé 逼近, 容易导出 $\exp(t\boldsymbol{B})$ 的矩阵逼近. 由定理 2.3.1 知

表 2.3.1 Padé 逼近表 (l,m)

$m \backslash l$	0	1	2	3	4
0	$\frac{1}{1}$	$\frac{1}{1-x}$	$\frac{1}{1-x+\frac{x^2}{2}}$	$\frac{1}{1-x+\frac{x^2}{2}-\frac{x^3}{6}}$	$\frac{1}{1-x+\frac{x^2}{2}-\frac{x^3}{6}+\frac{x^4}{24}}$
1	$\frac{1+x}{1}$	$\frac{1+\frac{x}{2}}{1-\frac{x}{2}}$	$\frac{1+\frac{x}{3}}{1-\frac{2x}{3}+\frac{x^3}{6}}$	$\frac{1+\frac{x}{4}}{1-\frac{3x}{4}+\frac{x^2}{4}-\frac{x^3}{24}}$	$\frac{1+\frac{x}{5}}{1-\frac{4x}{5}+\frac{3x^2}{10}-\frac{x^3}{15}+\frac{x^4}{120}}$
2	$\frac{1+x+\frac{x^2}{2}}{1}$	$\frac{1+\frac{2x}{3}+\frac{x^2}{6}}{1-\frac{x}{3}}$	$\frac{1+\frac{x}{2}+\frac{x^2}{12}}{1-\frac{x}{2}+\frac{x^2}{12}}$	$\frac{1+\frac{2x}{5}+\frac{x^2}{20}}{1-\frac{3x}{5}+\frac{x^2}{20}-\frac{x^3}{60}}$	$\frac{1+\frac{x}{3}+\frac{x^2}{30}}{1-\frac{2x}{3}+\frac{x^2}{20}-\frac{x^3}{30}+\frac{x^4}{360}}$
3	$\frac{1+x+\frac{x^2}{2}+\frac{x^3}{6}}{1}$	$\frac{1+\frac{3x}{4}+\frac{x^2}{4}+\frac{x^3}{24}}{1-\frac{x}{4}}$	$\frac{1+\frac{3x}{5}+\frac{3x^2}{20}+\frac{x^3}{60}}{1-\frac{2x}{5}+\frac{x^2}{20}}$	$\frac{1+\frac{x}{2}+\frac{x^2}{10}+\frac{x^3}{120}}{1-\frac{x}{2}+\frac{x^2}{10}-\frac{x^3}{120}}$	$\frac{1+\frac{3x}{7}+\frac{x^2}{14}+\frac{x^3}{210}}{1-\frac{4x}{7}+\frac{x^2}{7}-\frac{4x^3}{210}+\frac{x^4}{840}}$
4	$\frac{1+x+\frac{x^2}{2}+\frac{x^3}{6}+\frac{x^4}{24}}{1}$	$\frac{1+\frac{4x}{5}+\frac{3x^2}{10}+\frac{x^3}{15}+\frac{x^4}{120}}{1-\frac{x}{5}}$	$\frac{1+\frac{2x}{3}+\frac{x^2}{20}+\frac{x^3}{30}+\frac{x^4}{360}}{1-\frac{x}{3}+\frac{x^2}{30}}$	$\frac{1+\frac{4x}{7}+\frac{x^2}{7}+\frac{4x^3}{210}+\frac{x^4}{840}}{1-\frac{3x}{7}+\frac{x^2}{14}-\frac{x^3}{210}}$	$\frac{1+\frac{x}{2}+\frac{3x^2}{28}+\frac{x^3}{84}+\frac{x^4}{1680}}{1-\frac{x}{2}+\frac{3x^2}{28}-\frac{x^3}{84}+\frac{x^4}{1680}}$

道, 位于表 2.3.1 对角线上的有理式所构造的差分格式都应该是辛格式. 我们将表 2.3.1 中 $m=i, l=j$ 的元素记为 (i,j).

元素 $(1,1)$ (即 $l=1, m=1$) 对应 Euler 中点格式

$$\boldsymbol{z}^{k+1}=\boldsymbol{z}^k+\frac{\tau \boldsymbol{B}}{2}(\boldsymbol{z}^k+\boldsymbol{z}^{k+1}), \tag{2.3.8}$$

$$\boldsymbol{F}_\tau^{(1,1)}=\phi^{(1,1)}(\tau \boldsymbol{B}), \quad \phi^{(1,1)}(\lambda)=\frac{1+\dfrac{\lambda}{2}}{1-\dfrac{\lambda}{2}}, \tag{2.3.9}$$

此格式具有 2 阶精度.

元素 $(2,2)$ 对应的差分格式为

$$\boldsymbol{z}^{k+1}=\boldsymbol{z}^k+\frac{\tau \boldsymbol{B}}{2}(\boldsymbol{z}^k+\boldsymbol{z}^{k+1})+\frac{\tau^2 \boldsymbol{B}^2}{12}(\boldsymbol{z}^k-\boldsymbol{z}^{k+1}), \tag{2.3.10}$$

$$\boldsymbol{F}_\tau^{(2,2)}=\phi^{(2,2)}(\tau \boldsymbol{B}), \quad \phi^{(2,2)}(\lambda)=\frac{1+\dfrac{\lambda}{2}+\dfrac{\lambda^2}{12}}{1-\dfrac{\lambda}{2}+\dfrac{\lambda^2}{12}}, \tag{2.3.11}$$

此格式具有 4 阶精度.

元素 $(3,3)$ 对应差分格式

$$\begin{aligned}\boldsymbol{z}^{k+1}=&\boldsymbol{z}^k+\frac{\tau \boldsymbol{B}}{2}(\boldsymbol{z}^k+\boldsymbol{z}^{k+1})+\frac{\tau^2 \boldsymbol{B}^2}{10}(\boldsymbol{z}^k-\boldsymbol{z}^{k+1})\\&+\frac{\tau^3 \boldsymbol{B}^3}{120}(\boldsymbol{z}^k+\boldsymbol{z}^{k+1}),\end{aligned} \tag{2.3.12}$$

$$\boldsymbol{F}_\tau^{(3,3)}=\phi^{(3,3)}(\tau \boldsymbol{B}), \quad \phi^{(3,3)}(\lambda)=\frac{1+\dfrac{\lambda}{2}+\dfrac{\lambda^2}{10}+\dfrac{\lambda^3}{120}}{1-\dfrac{\lambda}{2}+\dfrac{\lambda^2}{10}-\dfrac{\lambda^3}{120}}, \tag{2.3.13}$$

此格式具有 6 阶精度.

元素 $(4,4)$ 对应差分格式

$$\begin{aligned}\boldsymbol{z}^{k+1}=&\boldsymbol{z}^k+\frac{\tau \boldsymbol{B}}{2}(\boldsymbol{z}^k+\boldsymbol{z}^{k+1})+\frac{3\tau^2 \boldsymbol{B}^2}{28}(\boldsymbol{z}^k-\boldsymbol{z}^{k+1})\\&+\frac{\tau^3 \boldsymbol{B}^3}{84}(\boldsymbol{z}^k+\boldsymbol{z}^{k+1})+\frac{\tau^4 \boldsymbol{B}^4}{1680}(\boldsymbol{z}^k-\boldsymbol{z}^{k+1}),\end{aligned} \tag{2.3.14}$$

$$F_\tau^{(4,4)} = \phi^{(4,4)}(\tau B),$$
$$\phi^{(4,4)}(\lambda) = \frac{1+\dfrac{\lambda}{2}+\dfrac{3\lambda^2}{28}+\dfrac{\lambda^3}{84}+\dfrac{\lambda^4}{1680}}{1-\dfrac{\lambda}{2}+\dfrac{3\lambda^2}{28}-\dfrac{\lambda^3}{84}+\dfrac{\lambda^4}{1680}}, \tag{2.3.15}$$

此格式具有 8 阶精度.

由 (2.3.4) 式和定理 2.3.1 可以证明下面的结论:

定理 2.3.2 线性系统 (2.2.1) 的差分格式

$$z^{k+1} = g_{ll}(\tau B)z^k, \quad l = 1, 2, \cdots$$

是 $2l$ 阶精度的辛格式.

§2.4 非线性 Hamilton 系统的辛差分格式

下面给出非线性 Hamilton 系统的几个常用的辛差分格式.

用中心差分格式离散方程 (2.0.1), 得到 Euler 中点格式[Fen85]

$$\frac{1}{\tau}(z^{k+1} - z^k) = J^{-1}H_z\left(\frac{z^{k+1}+z^k}{2}\right), \tag{2.4.1}$$

其步进映射 $F_\tau : z^k \mapsto z^{k+1}$ 是非线性的, 易求得它的偏导数

$$\frac{\partial z^{k+1}}{\partial z^k} = I + \tau J^{-1}H_{zz}\left(\frac{z^{k+1}+z^k}{2}\right)\left(\frac{1}{2}\frac{\partial z^{k+1}}{\partial z^k} + \frac{1}{2}I\right),$$

这里 $H_{zz}\left(\dfrac{z^{k+1}+z^k}{2}\right)$ 是函数 $H(z)$ 在点 $z = \dfrac{z^{k+1}+z^k}{2}$ 的 Hessian 矩阵. 整理得到步进映射的 Jacobi 矩阵 (记做 J_{F_τ}):

$$J_{F_\tau} = \left(I - \frac{\tau}{2}J^{-1}H_{zz}\left(\frac{z^{k+1}+z^k}{2}\right)\right)^{-1}\left(I + \frac{\tau}{2}J^{-1}H_{zz}\left(\frac{z^{k+1}+z^k}{2}\right)\right).$$

当 τ 充分小时, 无穷小阵 $\frac{\tau}{2}J^{-1}H_{zz}\left(\dfrac{z^{k+1}+z^k}{2}\right)$ 是非奇异的, 则 F_τ 是 Cayley 变换, 故 F_τ 是辛的.

值得注意的是, 和线性方程不一样, 非线性方程的首次积分 $\varphi(\boldsymbol{z})$ 包括 $H(\boldsymbol{z})$, 不一定能被以上格式精确保持, 它只满足近似守恒律:

$$\varphi(\boldsymbol{z}^{k+1})=\varphi(\boldsymbol{z}^{k})+o(\tau^3).$$

性质 2.4.1 若 $f(\boldsymbol{z})=\dfrac{1}{2}\boldsymbol{z}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{z}$ ($\boldsymbol{L}$ 为对称矩阵) 是 Hamilton 系统的一个二次守恒律, 则它也是 Euler 中点格式 (2.4.1) 的守恒律.

证明 由于

$$\begin{aligned}&\left\langle \boldsymbol{L}(\boldsymbol{z}^{k+1}+\boldsymbol{z}^{k}),\frac{\boldsymbol{z}^{k+1}-\boldsymbol{z}^{k}}{\tau}\right\rangle\\&\quad=\left\langle \boldsymbol{L}(\boldsymbol{z}^{k+1}+\boldsymbol{z}^{k}),\boldsymbol{J}^{-1}H_{\boldsymbol{z}}\left(\frac{\boldsymbol{z}^{k+1}+\boldsymbol{z}^{k}}{2}\right)\right\rangle\\&\quad=\left\langle (\boldsymbol{z}^{k+1}+\boldsymbol{z}^{k}),\boldsymbol{L}\boldsymbol{J}^{-1}H_{\boldsymbol{z}}\left(\frac{\boldsymbol{z}^{k+1}+\boldsymbol{z}^{k}}{2}\right)\right\rangle=0,\end{aligned}$$

其中 $\langle\cdot,\cdot\rangle$ 为标准内积, 最后一个等号是利用 $f(\boldsymbol{z})$ 为原 Hamilton 系统的守恒律得到的, 从而有 $\langle \boldsymbol{L}\boldsymbol{z}^{k},\boldsymbol{z}^{k}\rangle=\langle \boldsymbol{L}\boldsymbol{z}^{k+1},\boldsymbol{z}^{k+1}\rangle$. □

注 2.4.1 与 Euler 中点格式一样, 由 Padé 逼近表中对角线上其他的式子构造的辛格式也能保持原 Hamilton 系统的所有二次守恒律.

下面来看梯形格式

$$\frac{1}{\tau}(\boldsymbol{z}^{k+1}-\boldsymbol{z}^{k})=\boldsymbol{J}^{-1}\frac{1}{2}(H_{\boldsymbol{z}}(\boldsymbol{z}^{k+1})+H_{\boldsymbol{z}}(\boldsymbol{z}^{k})),\tag{2.4.2}$$

它的步进映射的 Jacobi 矩阵为

$$\boldsymbol{J}_{\boldsymbol{F}_\tau}=\left(\boldsymbol{I}-\frac{\tau}{2}\boldsymbol{J}^{-1}H_{\boldsymbol{zz}}(\boldsymbol{z}^{k+1})\right)^{-1}\left(\boldsymbol{I}+\frac{\tau}{2}\boldsymbol{J}^{-1}H_{\boldsymbol{zz}}(\boldsymbol{z}^{k})\right).$$

这个矩阵一般是非辛的, 因此以上的梯形格式是非辛的, 但是此格式可以通过一个变换修正成一个辛格式, 参见文献 [Dah75, QZZ95, WT03]. 引入非线性变换

$$\begin{aligned}\boldsymbol{\xi}^{k}&=\rho(\boldsymbol{z}^{k})=\boldsymbol{z}^{k}+\frac{\tau}{2}\boldsymbol{J}^{-1}H_{\boldsymbol{z}}(\boldsymbol{z}^{k}),\\\boldsymbol{\xi}^{k+1}&=\rho(\boldsymbol{z}^{k+1})=\boldsymbol{z}^{k+1}+\frac{\tau}{2}\boldsymbol{J}^{-1}H_{\boldsymbol{z}}(\boldsymbol{z}^{k+1}).\end{aligned}\tag{2.4.3}$$

上两等式相加, 得

$$\boldsymbol{\xi}^k + \boldsymbol{\xi}^{k+1} = \boldsymbol{z}^k + \boldsymbol{z}^{k+1} + \frac{\tau}{2}\boldsymbol{J}^{-1}\left(H_{\boldsymbol{z}}(\boldsymbol{z}^k) + H_{\boldsymbol{z}}(\boldsymbol{z}^{k+1})\right).$$

将梯形格式 (2.4.2) 代入上式, 得

$$\boldsymbol{\xi}^k + \boldsymbol{\xi}^{k+1} = \boldsymbol{z}^k + \boldsymbol{z}^{k+1} + \boldsymbol{z}^{k+1} - \boldsymbol{z}^k = 2\boldsymbol{z}^{k+1}.$$

把 $\boldsymbol{z}^{k+1} = \dfrac{\boldsymbol{\xi}^k + \boldsymbol{\xi}^{k+1}}{2}$ 代入 (2.4.3) 中的第二式, 得

$$\boldsymbol{\xi}^{k+1} = \frac{\boldsymbol{\xi}^k + \boldsymbol{\xi}^{k+1}}{2} + \frac{\tau}{2}\boldsymbol{J}^{-1}H_{\boldsymbol{z}}\left(\frac{\boldsymbol{\xi}^k + \boldsymbol{\xi}^{k+1}}{2}\right),$$

此即 Euler 中点格式

$$\boldsymbol{\xi}^{k+1} = \boldsymbol{\xi}^k + \tau\boldsymbol{J}^{-1}H_{\boldsymbol{z}}\left(\frac{\boldsymbol{\xi}^k + \boldsymbol{\xi}^{k+1}}{2}\right).$$

定理 2.4.1 梯形格式 (2.4.2) 保持如下辛结构[WT03]:

$$\boldsymbol{J} + \frac{\tau^2}{4}H_{\boldsymbol{zz}}(\boldsymbol{z})\boldsymbol{J}H_{\boldsymbol{zz}}(\boldsymbol{z}), \tag{2.4.4}$$

即

$$\begin{aligned}&\left(\frac{\partial \boldsymbol{z}^{k+1}}{\partial \boldsymbol{z}^k}\right)^{\mathrm{T}}\left(\boldsymbol{J} + \frac{\tau^2}{4}H_{\boldsymbol{zz}}(\boldsymbol{z}^{k+1})\boldsymbol{J}H_{\boldsymbol{zz}}(\boldsymbol{z}^{k+1})\right)\frac{\partial \boldsymbol{z}^{k+1}}{\partial \boldsymbol{z}^k}\\ &\quad = \boldsymbol{J} + \frac{\tau^2}{4}H_{\boldsymbol{zz}}(\boldsymbol{z}^k)\boldsymbol{J}H_{\boldsymbol{zz}}(\boldsymbol{z}^k).\end{aligned}$$

利用非线性变换 (2.4.3) 及命题 2.1.7 的结论, 直接计算即可证明定理 2.4.1.

注 2.4.2 对可分 Hamilton 系统, 即 $H(\boldsymbol{p},\boldsymbol{q}) = U(\boldsymbol{p}) + V(\boldsymbol{q})$, 可构造显式辛差分格式[Ruth83,Fen85]. 可分 Hamilton 方程和显式辛格式分别为

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = -V_{\boldsymbol{q}}(\boldsymbol{q}), \quad \frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}t} = U_{\boldsymbol{p}}(\boldsymbol{p}), \tag{2.4.5}$$

$$\begin{aligned}&\frac{1}{\tau}(\boldsymbol{p}^{k+1} - \boldsymbol{p}^k) = -V_{\boldsymbol{q}}(\boldsymbol{q}^{k+\frac{1}{2}}),\\ &\frac{1}{\tau}(\boldsymbol{q}^{k+1+\frac{1}{2}} - \boldsymbol{q}^{k+\frac{1}{2}}) = U_{\boldsymbol{p}}(\boldsymbol{p}^{k+1}).\end{aligned} \tag{2.4.6}$$

步进映射 $\boldsymbol{F}_\tau : \begin{pmatrix} \boldsymbol{p}^k \\ \boldsymbol{q}^{k+\frac{1}{2}} \end{pmatrix} \mapsto \begin{pmatrix} \boldsymbol{p}^{k+1} \\ \boldsymbol{q}^{k+1+\frac{1}{2}} \end{pmatrix}$ 的 Jacobi 矩阵形如

$$\boldsymbol{J}_{\boldsymbol{F}_\tau} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ -\tau U_{\boldsymbol{pp}} & \boldsymbol{I} \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{I} & -\tau V_{\boldsymbol{qq}} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix}.$$

由命题 2.1.5 知此 Jacobi 矩阵是辛阵.

性质 2.4.2 若 $f(\boldsymbol{p},\boldsymbol{q}) = \boldsymbol{p}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{q}$ 是方程 (2.4.5) 的守恒律, 则它的离散形式

$$(\boldsymbol{p}^{k+1})^{\mathrm{T}}\boldsymbol{L}\boldsymbol{q}^{k+\frac{3}{2}} = (\boldsymbol{p}^k)^{\mathrm{T}}\boldsymbol{L}\boldsymbol{q}^{k+\frac{1}{2}}$$

是差分方程 (2.4.6) 的守恒律.

证明 因为 $f(\boldsymbol{p},\boldsymbol{q})$ 是原 Hamilton 系统的守恒律, 所以有

$$\langle \boldsymbol{L}\boldsymbol{p}, U_{\boldsymbol{p}}(\boldsymbol{p})\rangle - \langle \boldsymbol{L}\boldsymbol{q}, V_{\boldsymbol{q}}(\boldsymbol{q})\rangle = 0.$$

又因为

$$\langle U_{\boldsymbol{p}}(\boldsymbol{p}^{k+1}), \boldsymbol{L}\boldsymbol{p}^{k+1}\rangle = \left\langle \frac{(\boldsymbol{q}^{k+\frac{3}{2}} - \boldsymbol{q}^{k+\frac{1}{2}})}{\tau}, \boldsymbol{L}\boldsymbol{p}^{k+1} \right\rangle,$$

$$\langle -V_{\boldsymbol{q}}(\boldsymbol{q}^{k+\frac{1}{2}}), \boldsymbol{L}\boldsymbol{q}^{k+\frac{1}{2}}\rangle = \left\langle \frac{(\boldsymbol{p}^{k+1} - \boldsymbol{p}^k)}{\tau}, \boldsymbol{L}\boldsymbol{q}^{k+\frac{1}{2}} \right\rangle,$$

整理得到

$$\langle \boldsymbol{L}\boldsymbol{p}^{k+1}, \boldsymbol{q}^{k+\frac{3}{2}}\rangle = \langle \boldsymbol{L}\boldsymbol{p}^k, \boldsymbol{q}^{k+\frac{1}{2}}\rangle. \qquad \square$$

§2.5 辛 R-K 方法及其相关方法

Runge-Kutta 方法 (简称 R-K 方法) 因为便于计算, 所以应用广泛. 但遗憾的是, 显式 R-K 方法通常不是辛格式. 不过, 对一类特殊的 Hamilton 系统——可分系统, 我们可以构造类似于 R-K 方法的辛格式, 而这些格式本质上是显式的.

2.5.1 多级辛 R-K 方法

考虑如下的 Hamilton 系统:

$$\begin{aligned}\frac{\mathrm{d}p_i}{\mathrm{d}t}&=-\frac{\partial H}{\partial q_i},\\ \frac{\mathrm{d}q_i}{\mathrm{d}t}&=\frac{\partial H}{\partial p_i},\end{aligned}\qquad i=1,\cdots,n, \tag{2.5.1}$$

其中 $H=H(p_1,\cdots,p_n,q_1,\cdots,q_n)$ 是 Hamilton 函数, 不显式依赖于时间 t. 系统 (2.5.1) 被称为**自治 Hamilton 系统**. 对于显式依赖于时间 t 的非自治系统, 可以通过引进两个新的自变量, 使它变成自治系统 (2.5.1) 的形式[Qin96].

首先将上面的方程改写成

$$\frac{\mathrm{d}\boldsymbol{z}}{\mathrm{d}t}=\boldsymbol{J}^{-1}H_{\boldsymbol{z}}=\boldsymbol{f}(\boldsymbol{z}). \tag{2.5.2}$$

方程 (2.5.2) 的单步 s 级 R-K 方法具有以下格式 (称为 R-K 格式):

$$\begin{aligned}\boldsymbol{z}^{k+1}&=\boldsymbol{z}^k+\tau\sum_{i=1}^{s}b_i\boldsymbol{f}(\boldsymbol{Y}_i),\\ \boldsymbol{Y}_i&=\boldsymbol{z}^k+\tau\sum_{j=1}^{s}a_{ij}\boldsymbol{f}(\boldsymbol{Y}_j),\quad i=1,\cdots,s,\end{aligned} \tag{2.5.3}$$

其中 $\tau=t_{k+1}-t_k\ (k=0,1,\cdots)$, $b_i,a_{ij}\ (i,j=1,\cdots,s)$ 是实的参数. 一个 R-K 格式的性质 (相容性、精度、稳定性等) 完全由这些参数所决定. 在 R-K 格式 (2.5.3) 中, 如果当 $j\geqslant i\ (i=1,\cdots,s)$ 时, $a_{ij}=0$, 则所有的 $\boldsymbol{Y}_i$ 可以通过 $\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_{i-1}$ 显式表示. 称这样的格式为**显式 R-K 格式**. 如果当 $j>i\ (i=1,\cdots,s-1)$ 时, $a_{ij}=0$, 并且在对角线上有某些 $a_{ii}\neq 0\ (1\leqslant i\leqslant s)$, 此时每一个 $\boldsymbol{Y}_i$ 都可以通过解一个 $2n$ 维的方程

$$\boldsymbol{Y}_i=\boldsymbol{z}^k+\tau\sum_{j=1}^{i-1}a_{ij}\boldsymbol{f}(\boldsymbol{Y}_j)+a_{ii}\boldsymbol{f}(\boldsymbol{Y}_i),\quad i=1,\cdots,s \tag{2.5.4}$$

得到, 称这种格式为**对角隐式 R-K 格式**. 一般隐式 R-K 格式既不是显式的, 也不是对角隐式的. 隐式 R-K 方法为了求 $\boldsymbol{Y}_i$ 每计算一步都

要求解一个 $s\times 2n$ 维的方程组, 当方程组是非线性时, 需要迭代求解, 导致不小的计算量. 因此, 在实际计算中, 人们往往偏爱显式 R-K 格式. 但隐式格式却有优于显式格式的稳定性, 所以我们不能忽视它.

1964 年, Butcher 提出一种表示格式 (2.5.3) 的简便方法[Butcher64], 称之为 **Butcher 向量法**, 即用如下向量表表示格式中的参数:

$$\begin{array}{c|c} \boldsymbol{c} & \boldsymbol{A} \\ \hline & \boldsymbol{b}^{\mathrm{T}} \end{array} \qquad \begin{array}{c|ccc} c_1 & a_{11} & \cdots & a_{1s} \\ c_2 & a_{21} & \cdots & a_{2s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{s1} & \cdots & a_{ss} \\ \hline & b_1 & \cdots & b_s \end{array} \tag{2.5.5}$$

其中 $c_i=\sum\limits_{j=1}^{s}a_{ij}\ (i=1,\cdots,s)$. 这样, 一个 R-K 格式可以由向量表 (2.5.5) 中的向量来确定. 向量表 (2.5.5) 称为 **Butcher 表**, 这种表达形式称为 **Butcher 表式**.

定义 2.5.1　如果格式 (2.5.3) 的步进映射是辛的, 即 Jacobi 矩阵 $\dfrac{\partial \boldsymbol{z}^{k+1}}{\partial \boldsymbol{z}^{k}}$ 是辛阵, 则称此格式是**辛 R-K 格式**.

定义 2.5.2　一个 s 级的 R-K 方法, **满足简化条件** $B(p),C(\eta)$, $D(\zeta)$ 分别是指如下条件满足:

$$\begin{aligned} &B(p):\boldsymbol{b}^{\mathrm{T}}\boldsymbol{c}^{k-1}=\frac{1}{k}, && 1\leqslant k\leqslant p,\\ &C(\eta):\boldsymbol{A}\boldsymbol{c}^{k-1}=\frac{1}{k}\boldsymbol{c}^{k}, && 1\leqslant k\leqslant \eta,\\ &D(\zeta):(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{c}^{k-1})^{\mathrm{T}}\boldsymbol{A}=\frac{1}{k}\big(\boldsymbol{b}^{\mathrm{T}}-(\boldsymbol{b}\boldsymbol{c}^{k})^{\mathrm{T}}\big), && 1\leqslant k\leqslant \zeta, \end{aligned}$$

其中 $\boldsymbol{A}$ 是 s 阶矩阵, $\boldsymbol{b}$ 和 $\boldsymbol{c}$ 是 s 维向量.

Butcher 在 1964 年证明了下列定理[Butcher64]:

定理 2.5.1　如果一个 R-K 格式的系数 $\boldsymbol{A},\boldsymbol{b},\boldsymbol{c}$, 满足简化条件 $B(p),C(\eta),D(\zeta)$ ($p\leqslant\eta+\zeta+1$, 且 $p\leqslant 2\eta+2$), 则该 R-K 格式有 p 阶精度.

记对角矩阵 $\mathrm{dig}(b_1,\cdots,b_s)$ 为 $\boldsymbol{B}, \boldsymbol{M}=\boldsymbol{BA}+\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}-\boldsymbol{b}\boldsymbol{b}^{\mathrm{T}}$. 下面我们介绍一个 R-K 格式为辛格式的充分条件. 这个条件首先由 Sanz-Serna 于 1988 年给出[SS88].

定理 2.5.2 如果 $\boldsymbol{M}=\boldsymbol{0}$, 则格式 (2.5.3) 是辛的.

注 2.5.1 特别地, 对于非约化的 R-K 格式 (若一个 s 级 R-K 方法不能等价降级为小于 s 级的 R-K 方法, 则称它为非约化的), $\boldsymbol{M}=\boldsymbol{0}$ 是辛 R-K 格式的充分必要条件.

下面介绍两个辛 R-K 格式: 1 级和 2 级 Gauss-Legendre 格式. 令 $c_i\,(i=1,\cdots,s)$ 为平移 Legendre 多项式 $Q_s(x)$ 的零点:

$$Q_s(x)=P_s\left(x-\frac{1}{2}\right), \tag{2.5.6}$$

$$P_s(x)=\frac{1}{2^s s!}\frac{\mathrm{d}^s}{\mathrm{d}x^s}\{(x^2-1)^s\}. \tag{2.5.7}$$

再令此 R-K 格式满足简化条件 $B(s), C(s)$, 于是从 $B(s)$ 即 $\sum\limits_{i=1}^{s} b_i c_i^{k-1}=\frac{1}{k}\ (k=1,\cdots,s)$ 可解出 $b_i\ (i=1,\cdots,s)$, 进而从 $C(s)$ 即 $\sum\limits_{j=1}^{s} a_{ij}c_j^{k-1}=\frac{1}{k}c_i^k\ (k,i=1,\cdots,s)$ 可解出 $a_{ij}\ (i,j=1,\cdots,s)$, 由此一个 R-K 格式得以确立. 通过这种方法构造的 R-K 格式称为 **Gauss-Legender 格式**.

我们给出 $s\leqslant 2$ 时 Gauss-Legendre 格式对应的 Butcher 表:

$s=1$:

$$\begin{array}{c|c} \frac{1}{2} & \frac{1}{2} \\ \hline & 1 \end{array} \tag{2.5.8}$$

$s=2$:

$$\begin{array}{c|cc} \frac{3-\sqrt{3}}{6} & \frac{1}{4} & \frac{3-2\sqrt{3}}{12} \\ \frac{3+\sqrt{3}}{6} & \frac{3+2\sqrt{3}}{12} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array} \tag{2.5.9}$$

可以看出, $s=1$ 时的格式就是我们所熟知的 Euler 中点格式. 而且不难验证, Butcher 表 (2.5.8) 和 (2.5.9) 均满足辛条件 $\boldsymbol{M}=\mathbf{0}$, 即它们都是辛 R-K 格式. 更进一步, 有以下结论:

定理 2.5.3 Gauss-Legendre 格式是具有 $2s$ 阶精度的辛格式.

证明 要证明是辛格式, 只需证明 $\boldsymbol{M}=\mathbf{0}$. 因为 Gauss-Legendre 格式满足简化条件 $D(s), C(s), B(2s)$, 所以有

$$\begin{aligned}\sum_{i=1}^{s} b_i a_{ij} c_i^{l-1} &= \frac{1}{l} b_j (1-c_j^l) = \frac{1}{l} b_j - \frac{1}{l} b_j c_j^l \\ &= \sum_{i=1}^{s} b_i b_j c_i^{l-1} - \sum_{i=1}^{s} b_j a_{ji} c_i^{l-1},\end{aligned}$$

即得

$$\sum_{i=1}^{s} (b_i a_{ij} + b_j a_{ji} - b_i b_j) c_i^{l-1} = 0, \quad l, j = 1, \cdots, s.$$

由于 $c_1, \cdots, c_s$ 互不相等, 因此 $\boldsymbol{M}=\mathbf{0}$. 由定理 2.5.2 知该格式是辛格式.

由定理 2.5.1 知, 该格式具有 $2s$ 阶精度. □

2.5.2 对角隐式辛 R-K 方法

在这一小节里, 我们给出一些对角隐式辛 R-K 格式, 这些辛格式具有计算方便、稳定性能好 (都是代数稳定的) 的优点, 参见文献 [QZ92, SSA91].

我们考虑满足条件 $\boldsymbol{M}=\mathbf{0}$ 的 s 级对角隐式 R-K 格式. 不失一般性, 假定 $b_i \neq 0$ $(i=1,\cdots,s)$. 令 $\boldsymbol{M}=\mathbf{0}$, 即有

$$b_i b_j - b_i a_{ij} - b_j a_{ji} = 0, \quad i, j = 1, \cdots, s. \tag{2.5.10}$$

如果 $b_k = 0$, 那么 $b_i a_{ik} = 0$ $(i=1,\cdots,s)$, 则此方法等价于一个级数小于 s 的方法, 即此方法是可约化的.

定理 2.5.4 如果 s 级对角隐式 R-K 格式满足条件 $\boldsymbol{M}=\mathbf{0}$, 则可

以将它表达成以下 Butcher 表:

$$
\begin{array}{c|cccccc}
c_1 & \frac{b_1}{2} \\
c_2 & b_1 & \frac{b_2}{2} \\
c_3 & b_1 & b_2 & \frac{b_3}{2} \\
\vdots & \vdots & \vdots & \vdots \\
c_s & b_1 & b_2 & b_3 & \cdots & \frac{b_s}{2} \\
\hline
 & b_1 & b_2 & b_3 & \cdots & b_s
\end{array}
\tag{2.5.11}
$$

其中 $c_i = \sum_{j=1}^{i} b_{j-1} + \frac{b_i}{2},\ i = 1, \cdots, s,\ b_0 = 0.$

证明 格式是对角隐式的, 因此 $a_{ij} = 0\ (j > i)$, 且满足 $\boldsymbol{M} = \boldsymbol{0}$, 即

$$b_i b_j - b_i a_{ij} - b_j a_{ji} = 0, \quad i, j = 1, 2, \cdots, s,$$

故

$$a_{ij} = b_j, \quad a_{ii} = \frac{b_i}{2}, \quad i = 1, \cdots, s,\ i > j.$$

□

推论 2.5.1 显式 R-K 格式不能满足条件 $\boldsymbol{M} = \boldsymbol{0}$.

注 2.5.2 Butcher 表 (2.5.11) 不能满足简化条件 $C(2)$.

下面我们给出 $s \leqslant 3$ 时的对角隐式辛 R-K 格式:

$s = 1$:

$$
\begin{array}{c|c}
\frac{1}{2} & \frac{1}{2} \\
\hline
 & 1
\end{array}
\tag{2.5.12}
$$

$s = 2$:

$$
\begin{array}{c|cc}
\frac{1}{4} & \frac{1}{4} & 0 \\
\frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\
\hline
 & \frac{1}{2} & \frac{1}{2}
\end{array}
\tag{2.5.13}
$$

$s = 3:$

$$
\begin{array}{c|ccc}
\frac{1}{2}a & \frac{1}{2}a & & \\
\frac{3}{2}a & a & \frac{1}{2}a & \\
\frac{1}{2}+a & a & a & \frac{1}{2}-a \\
\hline
 & a & a & 1-2a
\end{array}
\tag{2.5.14}
$$

其中 $a = 1.351207$, 它是多项式 $6x^3 - 12x^2 + 6x - 1$ 的实根[Cooper87]. 上面三个格式的精度分别为 $o(\tau^2), o(\tau^2), o(\tau^3)$.

推论 2.5.2 若 Butcher 表 (2.5.14) 的对角元素变为对称情况 $(a_{11} = a_{33})$, 即

$$
\begin{array}{c|ccc}
\frac{1}{2}a & \frac{1}{2}a & & \\
\frac{1}{2} & a & \frac{1}{2}-a & \\
1-\frac{1}{2}a & a & 1-2a & \frac{1}{2}a \\
\hline
 & a & 1-2a & a
\end{array}
\tag{2.5.15}
$$

则此格式达到 4 阶精度.

2.5.3 辛 P-R-K 方法

本小节介绍可分 Hamilton 系统的辛 R-K 方法. 假设 Hamilton 函数为

$$H(\boldsymbol{p}, \boldsymbol{q}) = U(\boldsymbol{p}) + V(\boldsymbol{q}), \tag{2.5.16}$$

对应可分 Hamilton 系统为

$$
\begin{aligned}
\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} &= -V_{\boldsymbol{q}}(\boldsymbol{q}) = \boldsymbol{f}(\boldsymbol{q}), \\
\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}t} &= U_{\boldsymbol{p}}(\boldsymbol{p}) = \boldsymbol{g}(\boldsymbol{p}).
\end{aligned}
\tag{2.5.17}
$$

若方程组 (2.5.17) 的第一个方程用一个 R-K 格式离散, 而第二个方程用另一个 R-K 格式离散, 这样的离散方法称为**分块 R-K 方法**, 即

P-R-K 方法, 对应两个 Butcher 表:

$$\begin{array}{c|ccc} c_1 & a_{11} & \cdots & a_{1s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{s1} & \cdots & a_{ss} \\ \hline & b_1 & \cdots & b_s \end{array} \qquad\qquad \begin{array}{c|ccc} C_1 & A_{11} & \cdots & A_{1s} \\ \vdots & \vdots & & \vdots \\ C_s & A_{s1} & \cdots & A_{ss} \\ \hline & B_1 & \cdots & B_s \end{array} \tag{2.5.18}$$

它们合起来构成 P-R-K 格式的 Butcher 表.

应用 Butcher 表 (2.5.18) 对方程组 (2.5.17) 进行计算, 得到

$$\begin{aligned} &\boldsymbol{P}_i = \boldsymbol{p}^n + \tau\sum_{j=1}^{s} a_{ij}\boldsymbol{f}(\boldsymbol{Q}_j), \\ &\qquad\qquad\qquad\qquad\qquad\qquad i = 1,\cdots,s, \\ &\boldsymbol{Q}_i = \boldsymbol{q}^n + \tau\sum_{j=1}^{s} A_{ij}\boldsymbol{f}(\boldsymbol{P}_j), \\ &\boldsymbol{p}^{n+1} = \boldsymbol{p}^n + \tau\sum_{i=1}^{s} b_i\boldsymbol{f}(\boldsymbol{Q}_i), \\ &\boldsymbol{q}^{n+1} = \boldsymbol{q}^n + \tau\sum_{i=1}^{s} B_i\boldsymbol{g}(\boldsymbol{P}_i). \end{aligned} \tag{2.5.19}$$

定理 2.5.5 如果 P-R-K 格式 (2.5.19) 的系数满足下面的条件:

$$\overline{M} \triangleq b_iA_{ij} + B_ja_{ji} - b_iB_j = 0, \quad i,j = 1,\cdots,s, \tag{2.5.20}$$

那么这个 P-R-K 格式是辛格式[AS93,SSM92,Sur90,Sun93].

2.5.4 辛 R-K-N 方法

下面考虑一类特殊的二阶常微分方程组, 这类方程组可以通过降阶变为 Hamilton 系统.

考虑二阶常微分方程

$$\ddot{\boldsymbol{y}} = \boldsymbol{f}(y_1,\cdots,y_n), \quad \boldsymbol{y} = (y_1,\cdots,y_n) \in \mathbf{R}^n, \tag{2.5.21}$$

其中 $\boldsymbol{f}(\boldsymbol{y})=(f_1(\boldsymbol{y}),\cdots,f_n(\boldsymbol{y}))^{\mathrm{T}}$ 是某一标量函数的梯度. 方程组 (2.5.21) 可以降阶成

$$\begin{pmatrix} \dot{\boldsymbol{y}} \\ \ddot{\boldsymbol{y}} \end{pmatrix} = \begin{pmatrix} \dot{\boldsymbol{y}} \\ \boldsymbol{f}(\boldsymbol{y}) \end{pmatrix}, \tag{2.5.22}$$

这里将 $\dot{y}_J$ $(J=1,\cdots,n)$ 视为一组新的未知变量. 因为 $\boldsymbol{f}(\boldsymbol{y})$ 是某一标量函数的梯度, 我们不妨假设 $\boldsymbol{f}(\boldsymbol{y})=\dfrac{\partial u}{\partial \boldsymbol{y}}$, 而 (2.5.22) 式可以改写为如下形式:

$$\begin{pmatrix}\dot{\boldsymbol{y}}\\ \ddot{\boldsymbol{y}}\end{pmatrix}=\begin{pmatrix}\dfrac{\partial H(\boldsymbol{y},\dot{\boldsymbol{y}})}{\partial \dot{\boldsymbol{y}}}\\ -\dfrac{\partial H(\boldsymbol{y},\dot{\boldsymbol{y}})}{\partial \boldsymbol{y}}\end{pmatrix}, \tag{2.5.23}$$

其中 $H(\boldsymbol{y},\dot{\boldsymbol{y}})$ 是标量函数, 满足

$$\frac{\partial H(\boldsymbol{y},\dot{\boldsymbol{y}})}{\partial \dot{\boldsymbol{y}}}=\dot{\boldsymbol{y}},$$
$$\frac{\partial H(\boldsymbol{y},\dot{\boldsymbol{y}})}{\partial \boldsymbol{y}}=-\boldsymbol{f}(\boldsymbol{y}).$$

这样的 Hamilton 函数 H 不唯一, 不妨令 $H=\dfrac{1}{2}\dot{\boldsymbol{y}}^{\mathrm{T}}\dot{\boldsymbol{y}}-u(\boldsymbol{y})$, $\dfrac{\partial u}{\partial \boldsymbol{y}}=\boldsymbol{f}(\boldsymbol{y})$. 系统 (2.5.23) 就是 Hamiltion 系统.

方程 (2.5.23) 如下形式的离散格式称为 **Runge-Kutta-Nyström 格式** (简称 **R-K-N 格式**):

$$\begin{aligned}
\boldsymbol{g}_i&=\boldsymbol{y}_0+c_i\tau\dot{\boldsymbol{y}}_0+\tau^2\sum_{j=1}^{s}a_{ij}\boldsymbol{f}(\boldsymbol{g}_j),\quad i=1,\cdots,s,\\
\boldsymbol{y}_1&=\boldsymbol{y}_0+\tau\dot{\boldsymbol{y}}_0+\tau^2\sum_{j=1}^{s}\overline{b}_j\boldsymbol{f}(\boldsymbol{g}_j),\\
\dot{\boldsymbol{y}}_1&=\dot{\boldsymbol{y}}_0+\tau\sum_{j=1}^{s}b_j\boldsymbol{f}(\boldsymbol{g}_j),
\end{aligned}\tag{2.5.24}$$

相应的离散方法称为 R-K-N 方法. 格式 (2.5.24) 的 Butcher 表为

$$\begin{array}{c|c}\boldsymbol{c} & \boldsymbol{A}\\ \hline & \overline{\boldsymbol{b}}^{\mathrm{T}}\\ & \boldsymbol{b}^{\mathrm{T}}\end{array}\qquad \text{即}\qquad \begin{array}{c|ccc}c_1 & a_{11} & \cdots & a_{1s}\\ c_2 & a_{21} & \cdots & a_{2s}\\ \vdots & \vdots & \cdots & \vdots\\ c_s & a_{s1} & \cdots & a_{ss}\\ \hline & \overline{b}_1 & \cdots & \overline{b}_s\\ & b_1 & \cdots & b_s\end{array}$$

定理 2.5.6 如果格式 (2.5.24) 的系数满足

$$\bar{b}_j = b_j(1 - c_j), \quad i = 1, \cdots, s, \tag{2.5.25}$$

$$b_i a_{ij} - b_j a_{ji} + \bar{b}_i b_j - b_i \bar{b}_j = 0, \quad i, j = 1, \cdots, s, \tag{2.5.26}$$

则它是辛格式.

定理 2.5.6 的证明参见文献 [Sur89]. 我们仅指出: 在条件 (2.5.25) 下, 条件 (2.5.26) 等价于

$$b_i a_{ij} - b_j a_{ji} + b_i b_j (c_j - c_i) = 0, \quad i, j = 1, \cdots, s. \tag{2.5.27}$$

参 考 文 献

[AS93] Abia L, Sanz-Serna J M. Partitioned Runge-Kutta methods for separable Hamiltonian problems. *Math Comp*, 1963, 60: 617–634.

[Butcher64] Butcher J C. Implicit Runge-Kutta Processes. *Math Comp*, 1964, 18: 50–64.

[Cooper87] Cooper G J. Stability of Runge-Kutta methods for trajectory problems. *IMA J Numer Anal*, 1987, 7: 1–13.

[Dah75] Dahlquist G. Error analysis for a class of methods for stiff non-linear initial value problems//Watson G A. *Numerical Analysis, Dundee*: *Lecture notes in Mathematics*. Berlin: Springer, 1975, 506: 60–74.

[Fen85] Feng K. On difference schemes and symplectic geometry//Feng K. *Proceedings of the 1984 Beijing Symposium on Differential Geometry and Differential Equations*. 北京: 科学出版社, 1985: 42–58.

[HNW93] Hairer E, Nørsett S P, Wanner G. *Solving Ordinary Differential Equations I, Nonstiff Problems*. 2nd edition. Berlin: Springer-Verlag, 1993.

[Las90] Lasagni F M. Integration methods for Hamiltonian differential equations. Unpublished manuscript, 1990.

[Qin96] Qin M Z. Symplectic difference schemes for nonautonomous Hamiltonian systemes. *Acta Applicandae Mathematicae*, 1996, 12(3): 309–321.

[QZ92] Qin M Z, Zhang M Q. Symplectic Runge-Kutta schemes for Hamiltonian system. *J Comput Math*: Supplementary Issue. 1992: 205–215.

[QZZ95] Qin M Z, Zhu W J, Zhang M Q. Construction of a three stage difference scheme for ordinary differential equations. *J Comput Math*, 1995, 13: 206–210.

[Ruth83] Ruth R. A canonical integration technique. *IEEE Trans Nucl Sci*, 1983, 30: 26–69.

[SS88] Sanz-Serna J M. Runge-Kutta schemes for Hamiltonian systems. *BIT*, 1988, 28: 877–883.

[SSA91] Sanz-Serna J M, Abia L. Order conditions for canonical Runge-Kutta schemes. *SIAM J Numer Anal*, 1991, 28: 1081–1096.

[SSM92] Saito S, Sugiura H, Mitsui T. Family of symplectic implicit Runge-Kutta formulae. *BIT*, 1992, 32: 539–543.

[Sun93] Sun G. Symplectic partitioned Runge-Kutta methods. *J Comput Math*, 1993, 11: 365–372.

[Sur89] Suris Y B. The canonicity of mappings generated by Runge-Kutta type methods when integrating the systems $\ddot{x} = -\frac{\partial U}{\partial x}$. *USSR Comput Maths Math Phys*, 1989, 29: 138–144.

[Sur90] Suris Y B. Hamiltonian methods of Runge-Kutta type and their variational interpretation. *Math Model*, 1990, 2: 78–87. (In Russian).

[WT03] Wang D L, Tam H W. A symplectic structure preserved by the trapezoidal rule. *J Phys Soc Japan*, 2003, 72(9): 2193–2197.

[FQ09] Feng K, Qin M Z. *Symplectic Geometric Algorithms for Hamiltonian Systems*. 浙江: Springer-Verlag GmbH, 2009.

[ZHW02] 张兴武, 武际可, 朱海平, 黄克复. 自治 Birkhoff 系统的广义正则变换和辛算法研究. 应用数学和力学, 2002, 23(9): 915–920.

[Mei96] 梅凤翔. 经典力学从牛顿到伯克霍夫. 力学与实践, 1996.

第三章　Euler-Lagrange 方程的变分对称性

给定一个偏微分方程，判定它是否为某个变分问题的 Euler-Lagrange 方程, 就是所谓的变分反问题. 这在有限元方法以及非线性波动理论等许多问题的研究中是一个很重要的组成部分. Hamilton 方程和 Birkhoff 方程就是由 Euler-Lagrange 方程变化而来的. 在历史上, 数学家应用泛函分析的方法对一般空间中算子位势性质的最早研究是 Vainberg 给出的判别准则, 但是正如文献 [AH75] 中指出的, 给出的结果很抽象, "不宜为大多数应用数学家和工程师们所了解". 所以, 直到 1973 年, Scott 等人在他们的一篇关于孤立子的有名的综述文章中还认为: 尚无一种一般的方法, 可据以证明一个方程没有 Lagrange 函数. 事实上, Tonti 在 1969 年就将 Vainberg 给出的结论运用到不同类型的微分方程, 引出了一些具体的判别法则[Ton69]. 这一工作为 Finlayson 所继续[Fin72a,Fin72b]. 1975 年, Atherton 和 Homsy 在上述工作的基础上, 运用泛函分析的方法得出了 n 阶非线性偏微分方程是 Euler-Lagrange 方程的一组充分必要条件[AH75]. 文献 [屠规彰等81] 又指出了文献 [AH75] 中多余的条件, 并就方程本身是未知函数及其各阶导数的多项式, 探讨了这类方程的结构. 在讨论一般方程的辛结构及多辛结构时, 判断方程决定的算子的位势性也是很必要的, 所以在这一章我们将从方程的对称性出发, 给出一套方便使用的方程及方程组的对称性判别准则, 并以此讨论多辛 Hamilton 方程和多辛 Birkhoff 方程的对称性.

§3.1　变分对称性

首先我们要介绍一些这里要用到的符号和基本概念. 为了方便, 我们在 Euclid 空间 $\mathbf{R}^n$ 表示的局部坐标空间上进行以下讨论, 局部空间上的结果也可以推广到光滑的微分流形上.

令 $X \subset \mathbf{R}^n$ 是一个有光滑边界 ∂X 的连通开集, M 表示一类定义在 $\mathbf{R}^n$ 上的光滑函数集合. 通过 Lagrange 函数 $L(u)$ 定义一类泛函

$$F(\boldsymbol{u}) = \int_X L(\boldsymbol{u})\mathrm{d}V, \tag{3.1.1}$$

其中 $\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{x}) \in M$, $\boldsymbol{x} \in X$, $L(\boldsymbol{u})$ 可以看做关于 $\boldsymbol{x}$, $\boldsymbol{u}$ 及 $\boldsymbol{u}$ 的导数的光滑函数, 在第八章就电磁场方程我们会给出其更具体的定义, 称之为泛函 $F(\boldsymbol{u})$ 的 **Lagrange 因子**. 一个泛函问题包括寻找泛函 $F(\boldsymbol{u})$ 在集合 M 上的临界点, 即要求 X 上的函数 $\boldsymbol{u}(\boldsymbol{x})$, 使得

$$\delta F(\boldsymbol{u}) = 0, \tag{3.1.2}$$

其中 δ 表示泛函变分求导数. 我们通过泛函的 Gâteaux 导数来定义变分导数. 设算子 $N : TM \to T^*M$, 对于 $\forall \boldsymbol{\phi} \in TM$ (我们不妨假设 M 是线性空间, 因此 $\boldsymbol{\phi}$ 也可以说是属于 M, 这样理解可以使后面的讨论更方便), 假使 $\boldsymbol{u} + \varepsilon\boldsymbol{\phi} \in M$ 始终成立, 那么我们定义算子 N 的 **Gâteaux 导数**为

$$N'_{\boldsymbol{u}}[\boldsymbol{\phi}] = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}(N(\boldsymbol{u} + \varepsilon\boldsymbol{\phi}) - N(\boldsymbol{u})), \tag{3.1.3}$$

如果 $N'_{\boldsymbol{u}}[\boldsymbol{\phi}]$ 关于 $\boldsymbol{\phi} \in M$ 总是连续的, 那么以上定义的也是 Fréchet 导数. 我们将 Fréchet 导算子 $N'_{\boldsymbol{u}}[\cdot]$ 简记为 $N'_{\boldsymbol{u}}$. 在以下讨论中, 我们假设 Gâteaux 导数总是连续的. 那么变分导数由下式导出:

$$\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}F(\boldsymbol{u} + \varepsilon\boldsymbol{\phi})\right|_{\varepsilon=0} = \int_X \delta F(\boldsymbol{u}) \cdot \boldsymbol{\phi}\mathrm{d}V, \quad \forall \boldsymbol{\phi} \in M, \varepsilon \in \mathbf{R}. \tag{3.1.4}$$

由 (3.1.3) 式和 (3.1.4) 式知

$$\int_X \delta F(\boldsymbol{u}) \cdot \boldsymbol{\phi}\mathrm{d}V = \int_X L'_{\boldsymbol{u}}[\boldsymbol{\phi}]\mathrm{d}V. \tag{3.1.5}$$

现在我们不妨假设变量 $u \in M$ 是一维的, 对于高维的情形, 可以做类似的讨论, 但我们在后面只给出高维情形的结果, 不再详细讨论. 利用分部积分法, 由上式推导出

$$\int_X \delta F(u) \cdot \phi \mathrm{d}V = \int_X E(L(u)) \cdot \phi \mathrm{d}V + \int_X \mathrm{div}(P_L)\mathrm{d}V, \tag{3.1.6}$$

其中 E 就是 Euler 算子:

$$E = \mathrm{D} + \sum_k \sum_{i_1,\cdots,i_k=1}^{n} (-1)^k \mathrm{D}_{x_{i_1}} \cdots \mathrm{D}_{x_{i_k}}, \tag{3.1.7}$$

$\mathrm{D} = \dfrac{\partial}{\partial u}$ 是关于 u 的偏导数, $\mathrm{D}_{x_{i_k}}$ 表示关于 x_{i_k} 的全导数, 即

$$\mathrm{D}_{x_{i_k}} = \frac{\partial}{\partial x_{i_k}} + u_{x_{i_k}} \frac{\partial}{\partial u} + \sum_{j=1}^{n} u_{x_j x_{i_k}} \frac{\partial}{\partial u_{x_j}} + \sum_{j,l=1}^{n} u_{x_j x_l x_{i_k}} \frac{\partial}{\partial u_{x_j x_l}} + \cdots, \tag{3.1.8}$$

而 $P_L(u,\phi)$ 是一个关于 $\boldsymbol{x}$, u 和 ϕ 以及它们的各阶导数的函数, 是对 (3.1.5) 式分部积分得到的边界项. 如果取边界条件 $P_L|_{\partial X} = 0$, 那么 $E(L(u)) = 0$ 就是泛函问题 $F(u) = \displaystyle\int_X L(u)\mathrm{d}V$ 对应的 Euler-Lagrange 方程.

下面给出几个有关方程对称性的定义.

定义 3.1.1 给出一个微分方程

$$N(u) = 0, \tag{3.1.9}$$

它对应的 Fréchet 导算子 N_u' 由下式给出:

$$N_u'[\phi] = \lim_{\varepsilon \to 0} \frac{N(u + \varepsilon\phi) - N(u)}{\varepsilon}, \quad \forall \phi \in M. \tag{3.1.10}$$

由上面定义的 Fréchet 导算子我们可以定义一个泛函:

$$\langle N_u'[\phi], \psi \rangle = \langle \psi, N_u'[\phi] \rangle = \int_X \psi N_u'[\phi] \mathrm{d}V, \tag{3.1.11}$$

其中 $\langle \cdot, \cdot \rangle$ 是 L^2 标准内积. 根据这个泛函我们又可以定义算子 N_u' 的共轭算子 $N_u'^*$.

定义 3.1.2 如果存在一个 Fréchet 导算子 B_u', 使得对 $\forall \psi, \phi \in M$, 有

$$\langle N_u'[\phi], \psi \rangle = \langle \phi, B_u'[\psi] \rangle + \int_X \mathrm{div}(P(u,\psi,\phi))\mathrm{d}V, \tag{3.1.12}$$

则称 B_u' 为 N_u' 的**共轭算子**, 记为 $N_u'^*$, 其中 P 为任意光滑函数.

定义 3.1.3 方程 (3.1.9) 对应的 Fréchet 导算子 N'_u 称为**自共轭算子**或**自伴算子**, 如果

$$N'_u = N'^*_u. \tag{3.1.13}$$

如果微分方程对应的 Fréchet 导算子 N'_u 是自共轭的, 则称这个微分方程是**对称**的.

我们是在相差一个散度函数的情形下定义这个共轭算子的, 因为在有关方程对称性的具体问题中, 这样做很必要. 在有些非线性波动理论或者泛函问题中, 边界条件的考虑是非常重要的. 一个泛函问题可以导出不同的微分方程和相应不同的边界条件. 相反地, 给出一个物理方程, 假如它是一个 Euler-Lagrange 方程, 如果附加一些合适的边界条件, 我们也可以导出不同的 Lagrange 泛函. 所以, 我们在定义共轭算子时, 强调的是其相差一个散度函数的情况下的唯一性. 事实上, 在多辛算法的具体例子中, 我们已经证实, 即使采取的边界条件不满足 $P|_{\partial X}=0$, 对计算结果也没有什么影响. 而在文献 [AH75] 中, 从基本的 Vainberg 定理出发, 给出的方程对称的充分必要条件中包含边界条件 $P|_{\partial X}=0$. 对于有些具体的应用, 这个条件过于严格, 比如波动方程

$$u_{tt}-u_{xx}=0, \tag{3.1.14}$$

它是 Euler-Lagrange 方程, 但是它显然不能满足 $\left.\dfrac{\partial N(u)}{\partial u_2}\right|_{\partial X}=0$, 这里 $(t,x)\in X$, $u_2=u_{tt},u_{xx}$.

在这里为了方便应用, 我们给出 Vainberg 定理的一个更一般推广:

定理 3.1.1 设算子 $N:TM\to T^*M$ 有 Fréchet 导算子 N'_u, 则 N 是一个梯度算子当且仅当对 $\forall u\in M$, N'_u 是自共轭算子; 梯度算子对应的微分方程 (3.1.9) 是以下变分问题的 Euler-Lagrange 方程:

$$F(u)=\int_X u\int_0^1 N(\lambda u)\mathrm{d}\lambda\mathrm{d}V. \tag{3.1.15}$$

证明 充分性 因为 N 有 Fréchet 导算子 N'_u, 所以可以计算 $F(u)$ 的变分导数, 得到

$$\int_X \delta F(u)\cdot\phi\mathrm{d}V=\int_X\left(\phi\int_0^1 N(\lambda u)\mathrm{d}\lambda+u\int_0^1 N'_{\lambda u}[\lambda\phi]\mathrm{d}\lambda\right)\mathrm{d}V. \tag{3.1.16}$$

又因为 N'_u 是自共轭算子, 即 $N'_u = N'^*_u$, 所以

$$\langle \psi, N'_u[\phi] \rangle = \langle \phi, N'^*_u[\psi] \rangle + \int_X \mathrm{div}(P(u,\psi,\phi))\mathrm{d}V, \tag{3.1.17}$$

从而有

$$\int_X \delta F(u) \cdot \phi \mathrm{d}V = \int_X \left(\phi \int_0^1 N(\lambda u)\mathrm{d}\lambda + \phi \int_0^1 N'^*_{\lambda u}[\lambda u]\mathrm{d}\lambda \right) \mathrm{d}V + \int_0^1 P(\lambda u, \lambda u, \phi)|_{\partial X} \mathrm{d}\lambda.$$

又由 $N'_u = N'^*_u$ 得

$$\int_X \delta F(u) \cdot \phi \mathrm{d}V = \int_X \left(\phi \int_0^1 N(\lambda u)\mathrm{d}\lambda + \phi \int_0^1 N'_{\lambda u}[\lambda u]\mathrm{d}\lambda \right) \mathrm{d}V + \int_0^1 P(\lambda u, \lambda u, \phi)\Big|_{\partial X} \mathrm{d}\lambda.$$

因为

$$N'_{\lambda u}[\lambda u] = \frac{\mathrm{d}}{\mathrm{d}\lambda}(\lambda N(\lambda u)) - N(\lambda u), \tag{3.1.18}$$

所以

$$\int_X \delta F(u) \cdot \phi \mathrm{d}V = \int_X \phi N(u)\mathrm{d}V + \int_0^1 P(\lambda u, \lambda u, \phi)\Big|_{\partial X} \mathrm{d}\lambda. \tag{3.1.19}$$

这说明 N 是 $F(u)$ 对应的梯度算子, 同时也说明方程 $N(u) = 0$ 是变分问题 (3.1.15) 的 Euler-Lagrange 方程.

必要性 充分性的证明过程可逆, 因此必要性也得证. □

§3.2 二阶 Euler-Lagrange 方程的对称性

我们给出一个任意阶的算子 N, 根据上一节 Fréchet 导算子及其

共轭算子的定义, 可给出直接计算 N_u' 和 $N_u'^*$ 的公式:

$$N_u' = \sum_{k=0}^{\infty} \sum_{i_1,\cdots,i_k=1}^{n} \frac{\partial N}{\partial u_{x_{i_1}\cdots x_{i_k}}} \mathrm{D}_{x_{i_1}} \cdots \mathrm{D}_{x_{i_k}}, \tag{3.2.1}$$

$$N_u'^* = \sum_{k=0}^{\infty} (-1)^k \sum_{i_1,\cdots,i_k=1}^{n} \mathrm{D}_{x_{i_1}} \cdots \mathrm{D}_{x_{i_k}} \left(\frac{\partial N}{\partial u_{x_{j_1}\cdots x_{j_s} x_{i_1}\cdots x_{i_k}}} \cdot \right) \mathrm{D}_{x_{i_1}} \cdots \mathrm{D}_{x_{i_k}}. \tag{3.2.2}$$

边界函数 P_N 如下:

$$P_N(u,\psi,\phi) = \sum_{k=1}^{\infty} \sum_{s=1}^{k} (-1)^{s-1} \sum_{i_1,\cdots,i_k=1}^{n} \mathrm{D}_{x_{i_1}} \cdots \mathrm{D}_{x_{i_{s-1}}} \left(\psi \frac{\partial N}{\partial u_{x_{i_1}\cdots x_{i_k}}} \right) \cdot \mathrm{D}_{x_{i_{s+1}}} \cdots \mathrm{D}_{x_{i_k}} \phi \bigg|_{\partial_{x_{i_s}} X}, \tag{3.2.3}$$

此处边界 $\partial X = \partial_{x_1} X \times \partial_{x_2} X \times \cdots \times \partial_{x_n} X$. 给定一个二阶偏微分方程

$$Lu = \sum_{i,j=1}^{n} a_{ij} u_{x_i x_j} + b = 0, \tag{3.2.4}$$

其中 a_{ij} 和 b 都是关于独立变量 x_i 和函数 u 及 u 的一阶导数 $u_{x_i} (i = 1, \cdots, n)$ 的光滑函数. 一个二阶偏微分方程要是 Euler-Lagrange 方程, 它首先应该是具有方程 (3.2.4) 的结构, 不可能再有其他形式. 这一点很容易证明, 这里不再赘述. ϕ 是 x_i 的光滑函数, 以它作为伴随变量, 我们得到 Lu 的 Fréchet 导数

$$L_u'[\phi] = \frac{\partial L}{\partial u}\phi + \sum_{k=1}^{n} \frac{\partial L}{\partial u_{x_k}} \mathrm{D}_{x_k}\phi + \sum_{i=1,j=1}^{n} \frac{\partial L}{\partial u_{x_i x_j}} \mathrm{D}_{x_i} \mathrm{D}_{x_j}\phi. \tag{3.2.5}$$

将 (3.2.4) 式代入 (3.2.5) 式, 得到

$$L_u'[\phi] = \sum_{i,j=1}^{n} a_{ij} \mathrm{D}_{x_i} \mathrm{D}_{x_j}\phi + \sum_{k=1}^{n} B \mathrm{D}_{x_k}\phi + A\phi, \tag{3.2.6}$$

其中

$$A = \sum_{i,j=1}^{n} u_{x_i x_j} \frac{\partial a_{ij}}{\partial u} + \frac{\partial b}{\partial u}, \quad B = \sum_{i,j=1}^{n} u_{x_i x_j} \frac{\partial a_{ij}}{\partial u_{x_k}} + \frac{\partial b}{\partial u_{x_k}}. \tag{3.2.7}$$

我们仍旧以 ϕ 作为伴随变量, 根据共轭算子的定义, 得到 L'_u 的共轭算子作用到 ϕ 上的表达式

$$\begin{aligned} L'^*_u[\phi] = & \sum_{i,j=1}^n a_{ij}\mathrm{D}_{x_i}\mathrm{D}_{x_j}\phi + \sum_{k=1}^n\left(\sum_{i=1}^n \mathrm{D}_{x_i}a_{ik} + \mathrm{D}_{x_i}a_{ki}\right)\mathrm{D}_{x_k}\phi \\ & + \sum_{i,j=1}^n(\mathrm{D}_{x_i}\mathrm{D}_{x_j}a_{ij})\phi - \sum_{k=1}^n B\mathrm{D}_{x_k}\phi - \sum_{k=1}^n(\mathrm{D}_{x_k}B)\phi + A\phi. \end{aligned} \tag{3.2.8}$$

根据上一节的定理 3.1.1, 方程 (3.2.4) 是 Euler-Lagrange 方程当下式成立:

$$L'_u - L'^*_u = 0. \tag{3.2.9}$$

将上式中各同阶项合并后, 令二阶项 $\mathrm{D}_{x_i}\mathrm{D}_{x_j}\phi$, 一阶项 $\mathrm{D}_{x_k}\phi$ 和 ϕ 的系数分别为零, 则得到

$$\begin{aligned} & a_{ij} = a_{ij}, \quad i,j=1,\cdots,n, \\ & B = \sum_{i=1}^n \mathrm{D}_{x_i}(a_{ik}+a_{ki}) - B, \\ & A = \sum_{j=1}^n \mathrm{D}_{x_i}\mathrm{D}_{x_j}a_{ij} - \sum_{k=1}^n \mathrm{D}_{x_k}B + A. \end{aligned} \tag{3.2.10}$$

经过计算, 上面的三个条件等价于下面两个等式:

$$B = \sum_{i=1}^n \mathrm{D}_{x_i}a_{ik}, \quad k=1,\cdots,n \tag{3.2.11}$$

和

$$\sum_{i,k=1}^n \mathrm{D}_{x_i}\mathrm{D}_{x_k}a_{ik} = \sum_{k=1}^n D_{x_k}B. \tag{3.2.12}$$

显然, 当 (3.2.11) 式成立时, (3.2.12) 式自然成立. 所以方程 (3.2.4) 是 Euler-Lagrange 方程的充分条件可以写为

$$\sum_{i=1}^n \mathrm{D}_{x_i}a_{ik} = \sum_{i,j=1}^n u_{x_ix_j}\frac{\partial a_{ij}}{\partial u_{x_k}} + \frac{\partial b}{\partial u_{x_k}}, \quad k=1,\cdots,n. \tag{3.2.13}$$

展开 D_{x_i}, 上式可以改写为

$$\sum_{i=1}^{n}\left(\frac{\partial a_{ik}}{\partial x_i}+\frac{\partial a_{ik}}{\partial u}u_{x_i}+\sum_{j=1}^{n}\frac{\partial a_{ik}}{\partial u_{x_j}}u_{x_ix_j}\right)=\sum_{i,j=1}^{n}u_{x_ix_j}\frac{\partial a_{ij}}{\partial u_{x_k}}+\frac{\partial b}{\partial u_{x_k}},\quad k=1,\cdots,n. \tag{3.2.14}$$

因为函数 a_{ij} 和 b 与 u 的二阶导数无关, 所以二阶导数项 $u_{x_ix_j}$ 除了在 $\sum\limits_{j=1}^{n}\frac{\partial a_{ik}}{\partial u_{x_j}}u_{x_ix_j}$ 和 $\sum\limits_{i,j=1}^{n}u_{x_ix_j}\frac{\partial a_{ij}}{\partial u_{x_k}}$ 中出现外, 在 (3.2.14) 式的其他项中不会出现. 令 $u_{x_ix_j}$ 的系数等于零, (3.2.14) 式变为

$$\begin{aligned}&\frac{\partial a_{ik}}{\partial u_{x_j}}=\frac{\partial a_{ij}}{\partial u_{x_k}},\\&\sum_{i=1}^{n}\left(\frac{\partial a_{ik}}{\partial x_i}+\frac{\partial a_{ik}}{\partial u}u_{x_i}\right)=\frac{\partial b}{\partial u_{x_k}},\end{aligned}\qquad k=1,\cdots,n. \tag{3.2.15}$$

以上就是一个二阶偏微分方程为 Euler-Lagrange 方程的最直接明了的充分条件.

推论 3.2.1　*一个二阶偏微分方程是某个变分问题的 Euler-Lagrange 方程, 当且仅当该方程形如方程 (3.2.4), 且该方程的系数满足 (3.2.15) 式.*

推论由定理 3.1.1 即可证.

§3.3　任意阶 Euler-Lagrange 方程组的对称性

我们在这一节引进微分算子矩阵和它的 Fréchet 导数以及共轭算子矩阵的概念, 用以描述偏微分方程组和它的对称性, 将上两节的结果做进一步的推广. 令 n 维微分流形 X 作为底空间, 有局部坐标 $\boldsymbol{x}=(x_1,\cdots,x_n)$, 目标微分流形有 l 维局部坐标 $\boldsymbol{u}=(u^1,\cdots,u^l)\in\mathbf{R}^l$. 令 $\mathfrak{X}=T^*_{\boldsymbol{x}}X\otimes T_{\boldsymbol{u}}\mathbf{R}^l$, 表示构形空间 $X\times\mathbf{R}^l$ 的切丛上的纤维, 假设它是一个实线性空间, $\mathfrak{X}^*$ 表示其共轭空间. 考虑一个 m 阶的由 l 个方程组成的偏微分系统

$$\boldsymbol{N}(\boldsymbol{x},\boldsymbol{u},\boldsymbol{u}(1),\cdots,\boldsymbol{u}(m))=\boldsymbol{0}, \tag{3.3.1}$$

其中 $\boldsymbol{u}(i)$ $(i=1,\cdots,m)$ 表示 $\boldsymbol{u}$ 的 i 阶混合导数, $\boldsymbol{N}=(N^1,\cdots,N^l)$ 是一组从构形空间的切丛纤维到微分算子的映射: $\mathfrak{X}\to\mathfrak{X}^*$, 不妨认为它就是一组算子向量, 它们的 Fréchet 导数是从 $\mathfrak{X}^*$ 到 $\mathfrak{X}$ 的算子构成的矩阵. 因为假设了 $\mathfrak{X}$ 是实线性的, 所以对 $\forall\boldsymbol{u}\in\mathfrak{X}$ 和 $\forall\boldsymbol{\phi}\in\mathfrak{X}^*$, 当 $\varepsilon\in\mathbf{R}$ 足够小时, 总有 $\boldsymbol{u}+\varepsilon\boldsymbol{\phi}\in\mathfrak{X}$.

定义 3.3.1 一个关于 n 个独立变量的最高为 m 阶的**微分算子**称为 m 阶 n 维**微分算子**; 由 l 个这样的算子组成的有序组称为一个 m 阶 $l\otimes n$ 维**微分算子向量**; 由 l 行 p 列这样的微分算子组成的表称为 m 阶 $l\times p\otimes n$ 维**微分算子矩阵**. 对微分算子矩阵中的每个元素求它的共轭算子得到的矩阵是原矩阵的**共轭微分算子矩阵**.

每给定一个微分方程组, 它的 Fréchet 导数总是对应着一个微分算子方阵. 我们用矩阵 $\boldsymbol{FD}$ 表示方程组 (3.3.1) 的 Fréchet 导数, 用 $\boldsymbol{FD}^*$ 表示它的共轭矩阵.

推论 3.3.1 方程组 (3.3.1) 的 Fréchet 导数是

$$\boldsymbol{FD}=\boldsymbol{N}'_{\boldsymbol{u}}=\sum_{k=0}^{m}\sum_{i_1,\cdots,i_k=1}^{n}\frac{\partial\boldsymbol{N}}{\partial u_{x_{i_1}\cdots x_{i_k}}}\mathrm{D}_{x_{i_1}}\cdots\mathrm{D}_{x_{i_k}},\tag{3.3.2}$$

它的共轭矩阵是

$$\boldsymbol{FD}^*=\boldsymbol{N}'^*_{\boldsymbol{u}}=\sum_{k=0}^{m}(-1)^k\sum_{i_1,\cdots,i_k=1}^{n}\mathrm{D}_{x_{i_1}}\cdots\mathrm{D}_{x_{i_k}}\left(\frac{\partial\boldsymbol{N}}{\partial u_{x_{i_1}\cdots x_{i_k}}}\cdot\right).\tag{3.3.3}$$

因为一个方程组就是一个向量函数的方程, 所以这个推论只需简单的计算就可以证明. 我们称 $\boldsymbol{FD}$ 是一个 m 阶 $l\times l\otimes n$ 维的微分算子方阵, 是指这个矩阵是 $l\times l$ 的, 它的元素所含自变量个数为 n, 矩阵中的元素是最高阶为 m 阶的导算子, 它的第 p 行 q 列元素就是方程组 (3.3.1) 中的第 p 个方程关于第 q 个未知函数求 Fréchet 导数, 即

$$\boldsymbol{FD}(p,q)=N^{p'}_{u^q}=\sum_{k=0}^{m}\sum_{i_1,\cdots,i_k=1}^{n}\frac{\partial N^p}{\partial u^q_{x_{i_1}\cdots x_{i_k}}}\mathrm{D}_{x_{i_1}}\cdots\mathrm{D}_{x_{i_k}},\quad p,q=1,\cdots,l.\tag{3.3.4}$$

推论 3.3.2 方程组 (3.3.1) 为对称的充分必要条件是它的 Fréchet 导数矩阵的转置等于它的共轭矩阵, 即

$$\boldsymbol{FD}^{\mathrm{T}} = \boldsymbol{FD}^*, \tag{3.3.5}$$

亦即
$$\boldsymbol{FD}^*(p,q) = \boldsymbol{FD}(q,p), \quad p,q = 1,\cdots,l.$$

推论由定理 3.1.1 即可证得.

§3.4 一阶 Euler-Lagrange 方程组的对称性

对于任何一个一阶微分方程来说, 根据本章第一节变分对称性的定义知道, 它一定不对称, 因此它也不是 Euler-Lagrange 方程. 对这种情况, 往往需要引进势函数, 将其变为二阶微分方程, 再讨论其对称性. 但是, 对于包含两个或者两个以上微分方程的一阶方程组来说, 可以直接讨论它的对称性, 不需要引进势函数.

设 $\boldsymbol{u} = (u^1,\cdots,u^l)$, $\boldsymbol{x} = (x_1,\cdots,x_n)$. 我们采用爱因斯坦记号习惯, 即每项中出现两次相同的角标要求和, 例如 $\dfrac{\partial F_j}{\partial z_i}\dfrac{\mathrm{d}z_j}{\mathrm{d}t} = \sum\limits_j \dfrac{\partial F_j}{\partial z_i}\dfrac{\mathrm{d}z_j}{\mathrm{d}t}$. 这样, 一阶微分方程组包括非线性方程组都可以表示成下面的形式:

$$\begin{gathered} N^p(u) = A^p_{\mu\nu}u^\nu_\mu + B^p = 0, \\ p,\nu = 1,\cdots,l, \quad \mu = 1,\cdots,n, \end{gathered} \tag{3.4.1}$$

其中 $B^p = B^p(\boldsymbol{x},\boldsymbol{u})$, $A^p_{\mu\nu} = A^p_{\mu\nu}(\boldsymbol{x},\boldsymbol{u},\boldsymbol{u_x})$, $u^\nu_\mu = \dfrac{\partial u^\nu}{\partial x_\mu}$. 我们求方程 (3.4.1) 的 Fréchet 导数算子和共轭导算子, 得到

$$\begin{aligned} N^{p'}_{u^q} &= \left(\frac{\partial A^p_{\mu\nu}}{\partial u^q}u^\nu_\mu + \frac{\partial B^p}{\partial u^q}\right) + A^p_{\mu q}\frac{\mathrm{d}}{\mathrm{d}x_\mu} + \frac{\partial A^p_{\mu\nu}}{\partial u^q_\beta}u^\nu_\mu\frac{\mathrm{d}}{\mathrm{d}x_\beta}, \\ N^{q'*}_{u^p} &= \left(\frac{\partial A^q_{\mu\nu}}{\partial u^p}u^\nu_\mu + \frac{\partial B^q}{\partial u^p}\right) - \frac{\mathrm{d}}{\mathrm{d}x_\mu}(A^q_{\mu p}) - A^q_{\mu p}\frac{\mathrm{d}}{\mathrm{d}x_\mu} \\ &\quad - \frac{\partial A^q_{\mu\nu}}{\partial u^p_\beta}u^\nu_\mu\frac{\mathrm{d}}{\mathrm{d}x_\beta} - \frac{\mathrm{d}}{\mathrm{d}x_\beta}\left(\frac{\partial A^q_{\mu\nu}}{\partial u^p_\beta}u^\nu_\mu\right). \end{aligned} \tag{3.4.2}$$

根据推论 3.3.3, 方程组对称当且仅当 $N_{u^q}^{p'} = N_{u^p}^{q'*}$. 我们发现, 假如 $\boldsymbol{A}_\mu = (A_{\mu\nu}^q)$ 中显含 $\boldsymbol{u}$ 的一阶导数, 那么在 $N_{u^p}^{q'*}$ 的最后一项中就会出现 $\boldsymbol{u}$ 的二阶导数. 而 $N_{u^q}^{p'}$ 中的系数最高只含 $\boldsymbol{u}$ 的一阶导数, 所以方程组对称的话, 系数 $\boldsymbol{A}_\mu$ 必定不显含 $\boldsymbol{u}$ 的一阶导数, 那么 $N_{u^q}^{p'}$ 和 $N_{u^p}^{q'*}$ 分别简化为

$$\begin{aligned} N_{u^q}^{p'} &= \left(\frac{\partial A_{\mu\nu}^p}{\partial u^q} u_\mu^\nu + \frac{\partial B^p}{\partial u^q}\right) + A_{\mu q}^p \frac{\mathrm{d}}{\mathrm{d}x_\mu}, \\ N_{u^p}^{q'*} &= \left(\frac{\partial A_{\mu\nu}^q}{\partial u^p} u_\mu^\nu + \frac{\partial B^q}{\partial u^p}\right) - \frac{\partial A_{\mu p}^q}{\partial u^\beta} u_\mu^\beta - \frac{\partial A_{\mu p}^q}{\partial x_\mu} - A_{\mu p}^q \frac{\mathrm{d}}{\mathrm{d}x_\mu}. \end{aligned} \tag{3.4.3}$$

令 $N_{u^q}^{p'} = N_{u^p}^{q'*}$, 得到下面三个条件:

$$A_{\mu q}^p = -A_{\mu p}^q, \tag{3.4.4}$$

$$\frac{\partial A_{\mu\nu}^p}{\partial u^q} u_\mu^\nu = \frac{\partial A_{\mu\nu}^q}{\partial u^p} u_\mu^\nu - \frac{\partial A_{\mu p}^q}{\partial u^\beta} u_\mu^\beta, \tag{3.4.5}$$

$$\frac{\partial B^q}{\partial u^p} - \frac{\partial B^p}{\partial u^q} = \frac{\partial A_{\mu p}^q}{\partial x_\mu}. \tag{3.4.6}$$

这是一个非常对称的等式结构. 对于每个固定的 $\mu = 1, \cdots, n$, 矩阵 $\boldsymbol{A}_\mu = (A_{\mu p}^q)$ 是反对称矩阵, 且满足 Jacobi 恒等式: 将 (3.4.5) 式中的 β 换为 ν, 再利用 (3.4.4) 式, 则 (3.4.5) 式就可以写为标准的 Jacobi 恒等式

$$\frac{\partial A_{\mu\nu}^p}{\partial u^q} + \frac{\partial A_{\mu q}^\nu}{\partial u^p} + \frac{\partial A_{\mu p}^q}{\partial u^\nu} = 0. \tag{3.4.7}$$

在第五章和第六章中, 我们将利用对称条件 (3.4.4), (3.4.5) 和 (3.4.6) 分别讨论辛 Birkhoff 系统和多辛 Birkhoff 系统的对称性.

方程组 (3.4.1) 可以紧凑地表示为

$$\boldsymbol{N}(\boldsymbol{u}) = \boldsymbol{A}_\mu \boldsymbol{u}_{x_\mu} + \boldsymbol{B} = \boldsymbol{0}, \tag{3.4.8}$$

其中 $\boldsymbol{A}_\mu$ 代表沿 x_μ 方向的辛结构. (3.4.6) 式可以改写为

$$\frac{\partial B^q}{\partial u^p} - \frac{\partial B^p}{\partial u^q} = \frac{\partial A_1(q,p)}{\partial x_1} + \cdots + \frac{\partial A_n(q,p)}{\partial x_n}, \quad p, q = 1, \cdots, l. \tag{3.4.9}$$

假如 (3.4.9) 式不为零, 表示函数向量 $\boldsymbol{B}$ 本身不是对称的, 即它不是某个标量势函数的梯度函数. 而沿着各个自变量方向的辛结构 $\boldsymbol{A}_\mu$ 产生的耗散项和函数 $\boldsymbol{B}$ 所产生的耗散平衡了, 所以把辛结构和 $\boldsymbol{B}$ 函数合起来, 整个系统依然是对称的, 这样的方程组就是 Birkhoff 系统.

如果 (3.4.9) 式等于零, 假如方程是常微分的, 那么它就是辛结构不依赖于自变量的 Hamilton 系统. 对于偏微分方程, 有两种情况: 一是, 各个方向的辛结构不显含本方向的自变量; 二是, 各个辛结构产生的耗散恰好互相平衡掉了, 这样方程分为两部分, 方程除了整体对称外, 带有辛结构的部分和 Hamilton 函数部分还各自对称.

参 考 文 献

[屠规彰等 81] 屠规彰，秦孟兆. Euler-Lagrange 方程的判别准则. 数学学报, 1981, 24(2): 190–206.

[AH75] Atherton R W, Homsy G M. On the existence and formulation of variational principles for nonlinear differential equations. *Stud Appl Math*, 1975, LIV (1): 1531–1551.

[Fin72a] Finlayson B A. Existence of variational principles for the Navier-Stokes equation. *Phys Fluids*, 1972, 15: 963–967.

[Fin72b] Finlayson B A. *The Method of Weighted Residuals and Variational Principles.* New York: Academic Press, 1972.

[Ton69] Tonti E. Variational formulations of nonlinear differential equations I, II. *Acad Roy Belg Bull Cl Sci*, 1969, 55: 137–165, 262–278.

第四章　受外力作用的系统的几何变分方法

受外力作用的 Lagrange 系统和 Hamilton 系统存在于各种现实的物理环境中, 比如受到激励器械施加力、磨擦力以及受到负荷的机械系统. 考虑这些系统的几何结构, 也是避免在数值计算中走弯路的一个好办法. 文献 [MW01, Arn63, Mos62] 分别讨论了一类受外力作用的 Lagrange 系统和 Hamilton 系统 (包括约束 Hamilton 系统) 的离散变分原理, 给出了构造这类系统保结构的变分积分子的几何变分方法.

基于几何变分, 进而可以构造 Veselov 型的离散变分原理. 而对于常微分的 Hamilton 系统, Veselov 型离散变分可以导出具有很多良好性质的辛积分子. Veselov 于 1988 年和 1991 年相继提出和发展了 Veselov 离散力学[Ves88, Ves91, FQ91] (它是基于离散的 Hamilton 原理发展的一种离散方法). 但是以上所说的几何变分没有包括一类特殊的变分微分方程. 这类特殊的变分微分方程的变分形式是 Pfaff 作用泛函形式, 它们的 Lagrange 函数是奇异的, 即 $\dfrac{\partial^2 L}{\partial \dot{q}^i \dot{q}^j} = 0$, 而且它们包括了受外力作用的系统. 对于这类方程, 我们利用几何变分的思想, 在不同的几何框架下对 Pfaff 作用泛函做三种不同形式的变分: 一般相空间的变分、增广空间的变分以及包含区域变分的全变分, 并给出对应的 Veselov 离散形式. 比较这三种变分, 我们最后得到的 Euler-Lagrange 方程是完全一致的. 但是和第一种变分不同的是, 后面两种变分得到了系统的一组形式能量和动量公式, 它们也可以看做 Euler-Lagrange 方程附加的约束条件. 第一种变分直接导出辛流公式, 而第二种变分可以导出 Lagrange 流在包含时间变量的增广空间上保持一个几何结构, 在标准相空间上保持另外一个辛结构. 第三种变分即全变分在包含时间 t 和 Hamilton 量的增广空间上也是保辛的.

在本章第一节中, 我们首先回顾一下 Marsden 几何变分方法的

Veselov 离散[MW01]. 在后面几节中, 我们将这种几何变分的思想加以推广, 用来讨论 Pfaff 作用泛函问题的几何变分结构.

§4.1 受外力作用的几何变分

4.1.1 受外力作用的 Lagrange 系统和 Hamilton 系统

定义 4.1.1 保纤维映射

$$f : TM \to T^*M,$$
$$(\boldsymbol{q}, \dot{\boldsymbol{q}}) \mapsto (\boldsymbol{q}, f(\boldsymbol{q}, \dot{\boldsymbol{q}}))$$

称为一个 **Lagrange 力**. 这时通常将 Lagrange 力记为 f_L.

给定一个 Lagrange 泛函

$$\mathcal{L} = \int_0^T L(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t))\mathrm{d}t, \tag{4.1.1}$$

其中 $L(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t))$是 Lagrange 函数, 要求曲线 $\boldsymbol{q} \in M$, 使得

$$\delta \int_0^T L(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t))\mathrm{d}t + \int_0^T f_L(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t)) \cdot \delta \boldsymbol{q}(t)\mathrm{d}t = 0. \tag{4.1.2}$$

利用分部积分法, 我们得到受外力作用的 Euler-Lagrange 方程

$$\frac{\partial}{\partial \boldsymbol{q}} L(\boldsymbol{q}, \dot{\boldsymbol{q}}) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial}{\partial \dot{\boldsymbol{q}}} L(\boldsymbol{q}, \dot{\boldsymbol{q}})\right) + f_L(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \boldsymbol{0}. \tag{4.1.3}$$

这个方程就是在普通 Euler-Lagrange 方程上加了一个外力项.

定义 4.1.2 保纤维映射 $f_H : T^*M \to T^*M$ 称为 **Hamilton 力**.

给定一个 Hamilton 力 f_H, 我们在 T^*M 上定义一个对应的 1-形式 f'_H:

$$f'_H(\boldsymbol{p_q})(\boldsymbol{u_{p_q}}) = \langle f_H(\boldsymbol{p_q}), T_{\pi_M}(\boldsymbol{u_{p_q}})\rangle, \tag{4.1.4}$$

其中 $\boldsymbol{p_q} \in T^*M, \boldsymbol{u_{p_q}} \in T(T^*M)$ (之所以用 $\boldsymbol{p_q}$ 表示 T^*M 中的元素, 是因为 $(\boldsymbol{q}, \dot{\boldsymbol{q}})$ 构成 TM, 而 $\boldsymbol{p} = \dot{\boldsymbol{q}}$), $\pi_M : T^*M \to M$ 是投影影射. f'_H 和 T^*M 上的典则 2-形式 Θ 类似:

$$\Theta(\boldsymbol{p_q})(\boldsymbol{u_{p_q}}) = \langle \boldsymbol{p_q}, T\pi_{T^*M}(\boldsymbol{u_{p_q}})\rangle, \tag{4.1.5}$$

其中 $\pi_{T^*M}: T(T^*M) \to M$ 是 π_M 的提升.

f'_H 在局部坐标上表示为

$$f'_H(\boldsymbol{q},\boldsymbol{p}) \cdot (\delta\boldsymbol{q},\delta\boldsymbol{p}) = f_H(\boldsymbol{q},\boldsymbol{p}) \cdot \delta\boldsymbol{q}, \tag{4.1.6}$$

所以形式 f'_H 是一个水平形式. 受外力作用的 Hamilton 向量场 X_H 由下式定义:

$$i_{X_H}\omega = \mathrm{d}H - f'_H, \tag{4.1.7}$$

其中 ω 为标准 Hamilton 辛形式. 公式 (4.1.7) 在坐标意义下给出了著名的受外力作用的 Hamilton 方程组:

$$\begin{aligned} X_{\boldsymbol{q}}(\boldsymbol{q},\boldsymbol{p}) &= \frac{\partial}{\partial\boldsymbol{q}}H(\boldsymbol{q},\boldsymbol{p}), \\ X_{\boldsymbol{p}}(\boldsymbol{q},\boldsymbol{p}) &= \frac{\partial}{\partial\boldsymbol{p}}H(\boldsymbol{q},\boldsymbol{p}) + f_H(\boldsymbol{q},\boldsymbol{p}), \end{aligned} \tag{4.1.8}$$

它和标准 Hamilton 系统的区别在于动量方程上增加了外力项.

4.1.2 受外力作用的 Legendre 变换和 Noether 定理

给定一个 Lagrange 函数 L, 通过标准的 Legendre 变换 $\mathbb{F}L: T^*M \to TM$, 将其和 Hamilton 函数联系起来:

$$H(\boldsymbol{q},\boldsymbol{p}) = \mathbb{F}L(\boldsymbol{q},\dot{\boldsymbol{q}}) \cdot \dot{\boldsymbol{q}} - L(\boldsymbol{q},\dot{\boldsymbol{q}}). \tag{4.1.9}$$

Lagrange 力 f_L 和 Hamilton 力 f_H 也通过 Legendre 变换联系起来了:

$$f_L = \mathbb{F}L \circ f_H. \tag{4.1.10}$$

那么就可以证明受外力作用的 Euler-Lagrange 方程 (4.1.3) 和受外力作用的 Hamilton 方程组 (4.1.8) 是等价的. 如果用 X_L 和 X_H 分别表示受外力作用的 Lagrange 向量场和 Hamilton 向量场, 则有

$$(\mathbb{F}L)^*(X_H) = X_L.$$

假设存在一个李群 G, 它的李代数为 $\mathcal{G}$, 通过作用函数 $\boldsymbol{\Phi}$: $G \times M \to M$ 作用在 M 上. 考虑对应的作用 $\boldsymbol{\Phi}^{TM}: G \times TM \to TM$,

定义为 $\boldsymbol{\Phi}^{TM}(\boldsymbol{g},\boldsymbol{v_q}) = T(\boldsymbol{\Phi})\cdot \boldsymbol{v_q}$, $\forall \boldsymbol{g}\in G$, $\boldsymbol{v_q}\in TM$, 在局部坐标下表示为

$$\boldsymbol{\Phi}^{TM}(\boldsymbol{g},(\boldsymbol{q},\dot{\boldsymbol{q}})) = \left(\Phi^i(\boldsymbol{g},\boldsymbol{q}), \frac{\partial}{\partial q_j}\Phi^i(\boldsymbol{g},\boldsymbol{q})\dot{q}_j\right). \tag{4.1.11}$$

设 $\boldsymbol{\xi}\in\mathcal{G}$, 无穷小生成子 $\boldsymbol{\xi}_M: M\to TM$ 和 $\boldsymbol{\xi}_{TM}: TM\to T(TM)$ 定义为

$$\boldsymbol{\xi}_M(\boldsymbol{q}) = \frac{\mathrm{d}}{\mathrm{d}\boldsymbol{g}}(\boldsymbol{\Phi}(\boldsymbol{g},\boldsymbol{q}))\cdot\boldsymbol{\xi},\quad \boldsymbol{\xi}_{TM}(\boldsymbol{v_q}) = \frac{\mathrm{d}}{\mathrm{d}\boldsymbol{g}}(\boldsymbol{\Phi}^{TM}(\boldsymbol{g},\boldsymbol{v_q}))\cdot\boldsymbol{\xi}. \tag{4.1.12}$$

TM 的局部坐标为 $(\boldsymbol{q},\dot{\boldsymbol{q}})$, $T(TM)$ 的局部坐标为 $((\boldsymbol{q},\dot{\boldsymbol{q}}),(\dot{\boldsymbol{q}},\ddot{\boldsymbol{q}}))$. 在坐标意义下, 无穷小生成子表示为

$$\begin{aligned}\boldsymbol{\xi}_M(\boldsymbol{q}) &= \left(q_i, \frac{\partial}{\partial g^m}\Phi^i(\boldsymbol{e},\boldsymbol{q})\xi^m\right),\\ \boldsymbol{\xi}_{TM}(\boldsymbol{q},\dot{\boldsymbol{q}}) &= \left((q_i,\dot{q}_i), \left(\frac{\partial}{\partial g^m}\Phi^i(\boldsymbol{e},\boldsymbol{q})\xi^m, \frac{\partial^2}{\partial g^m\partial q_j}\Phi^i(\boldsymbol{e},\boldsymbol{q})\dot{q}_j\xi^m\right)\right),\end{aligned} \tag{4.1.13}$$

其中 $\boldsymbol{e}$ 表示 G 的单位元. 定义映射 $J_L: TM\to\mathcal{G}^*$ 为 **Lagrange 动量映射**, 使得

$$J_L(\boldsymbol{v_q})\cdot\boldsymbol{\xi} = \Theta_L\cdot\boldsymbol{\xi}_{TM}(\boldsymbol{v_q}). \tag{4.1.14}$$

其中 Θ_L 称为 **Lagrange 1-形式**:

$$\Theta_L\cdot\boldsymbol{\xi} = \left\langle\frac{\partial L}{\partial\dot{\boldsymbol{q}}}, \boldsymbol{\xi}_M(\boldsymbol{q})\right\rangle, \tag{4.1.15}$$

这里 $\dfrac{\partial L}{\partial\dot{\boldsymbol{q}}}$ 代表 Legendre 变换. 令 $\delta\boldsymbol{q}(t) = \boldsymbol{\xi}_M(\boldsymbol{q}(t))$, 求变分 (4.1.2), 得到

$$\int_0^T \mathrm{d}L\cdot\boldsymbol{\xi}_{TM}\mathrm{d}t + \int_0^T f_L\cdot\boldsymbol{\xi}_M\mathrm{d}t = 0. \tag{4.1.16}$$

假设 L 是不变的, F_L^t 代表其相流, 对上式进行分部积分, 可推导出

$$\begin{aligned}&\int_0^T\left[\frac{\partial}{\partial\boldsymbol{q}}L(\boldsymbol{q},\dot{\boldsymbol{q}}) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial}{\partial\dot{\boldsymbol{q}}}L(\boldsymbol{q},\dot{\boldsymbol{q}})\right) + f_L(\boldsymbol{q},\dot{\boldsymbol{q}})\right]\cdot\boldsymbol{\xi}_M\mathrm{d}t + \Theta_L\cdot\boldsymbol{\xi}_{TM}\Big|_0^T\\ &= (J_L\circ F_L^T)(\boldsymbol{q}(0),\dot{\boldsymbol{q}}(0))\cdot\boldsymbol{\xi} - J_L(\boldsymbol{q}(0),\dot{\boldsymbol{q}}(0))\cdot\boldsymbol{\xi}.\end{aligned} \tag{4.1.17}$$

比较 (4.1.17) 式和 (4.1.16) 式, 得到

$$
\begin{aligned}
&[(J_L \circ F_L^T)(\boldsymbol{q}(0), \dot{\boldsymbol{q}}(0)) - J_L(\boldsymbol{q}(0), \dot{\boldsymbol{q}}(0))] \cdot \boldsymbol{\xi} \\
&= \int_0^T f_L(\boldsymbol{q}(t), \dot{\boldsymbol{q}}(t)) \cdot \boldsymbol{\xi}_M(\boldsymbol{q}(t)) \mathrm{d}t. \qquad (4.1.18)
\end{aligned}
$$

上面的等式说明了外力是如何影响 Lagrange 动量映射从初始时刻 0 到时刻 T 的变化的. 如果外力作用是垂直于前面定义的群作用的, 即 $f_L(\boldsymbol{q}, \dot{\boldsymbol{q}}) \cdot \boldsymbol{\xi}_M(\boldsymbol{q}) = 0$, 在这种特殊情况下, 我们得到如下 Noether 定理:

定理 4.1.1 假设一个 Lagrange 系统 $L : TM \to \mathbf{R}$ 受到一个 Lagrange 力 $f_L : TM \to T^*M$ 的作用, 而且存在对称作用 $\boldsymbol{\Phi} : G \times M \to M$, 使得对所有的 $(\boldsymbol{q}, \dot{\boldsymbol{q}}) \in TM$ 和所有的 $\boldsymbol{\xi} \in \mathcal{G}$, 有 $\langle f_L(\boldsymbol{q}, \dot{\boldsymbol{q}}), \boldsymbol{\xi}_M(q) \rangle = 0$, 那么变分相流 F_L^t 保持 Lagrange 动量映射 $J_L : TM \to \mathcal{G}^*$, 即对任意 t, 有

$$J_L \circ F_L^t = J_L.$$

对于受外力作用的 Hamilton 系统, 通过对 Lagrange 系统作 Legendre 变换可以得到类似的结果, 或者直接从 Hamilton 系统 (4.1.8) 出发可以推出类似的结论. 值得注意的是, 受外力作用的 Lagrange 流和 Hamilton 相流并不保持辛形式.

4.1.3 受外力作用的离散变分力学

首先我们回顾一下 Marsden 几何变分方法的 Veselov 离散 [MW01].

用 $M \times M$ 作为构形空间的切丛 TM 的离散空间. 假定时间步长为 τ, 则 $M \times M$ 中的某一点 $(\boldsymbol{q}_1, \boldsymbol{q}_0)$ 对应切向量 $\dfrac{\boldsymbol{q}_1 - \boldsymbol{q}_0}{\tau}$. 定义一个光滑映射

$$L : M \times M = \{(\boldsymbol{q}_1, \boldsymbol{q}_0)\} \to \mathbf{R},$$

其离散的作用泛函为

$$\mathcal{L} = \sum_{k=1}^{N} L(\boldsymbol{q}_k, \boldsymbol{q}_{k-1}). \qquad (4.1.19)$$

离散变分原理就是求 $(\boldsymbol{q}_0, \cdots, \boldsymbol{q}_N)$, 使得 $\mathcal{L}$ 取极值. 对上面的离散泛函变分, 得到

$$\begin{aligned}
&\mathrm{d}\mathcal{L}(\boldsymbol{q}_0, \cdots, \boldsymbol{q}_N)(\delta\boldsymbol{q}_0, \cdots, \delta\boldsymbol{q}_N) \\
&= \sum_{k=0}^{N-1}\left[\frac{\partial}{\partial\boldsymbol{q}_1}L(\boldsymbol{q}_{k+1}, \boldsymbol{q}_k)\delta\boldsymbol{q}_{k+1} + \frac{\partial}{\partial\boldsymbol{q}_0}L(\boldsymbol{q}_{k+1}, \boldsymbol{q}_k)\delta\boldsymbol{q}_k\right] \\
&= \sum_{k=1}^{N}\frac{\partial}{\partial\boldsymbol{q}_1}L(\boldsymbol{q}_k, \boldsymbol{q}_{k-1})\delta\boldsymbol{q}_k + \sum_{k=0}^{N-1}\frac{\partial}{\partial\boldsymbol{q}_0}L(\boldsymbol{q}_{k+1}, \boldsymbol{q}_k)\delta\boldsymbol{q}_k \\
&= \sum_{k=1}^{N-1}\left[\frac{\partial}{\partial\boldsymbol{q}_1}L(\boldsymbol{q}_k, \boldsymbol{q}_{k-1}) + \frac{\partial}{\partial\boldsymbol{q}_0}L(\boldsymbol{q}_{k+1}, \boldsymbol{q}_k)\right]\delta\boldsymbol{q}_k \\
&\quad + \frac{\partial}{\partial\boldsymbol{q}_0}L(\boldsymbol{q}_1, \boldsymbol{q}_0)\delta\boldsymbol{q}_0 + \frac{\partial}{\partial\boldsymbol{q}_1}L(\boldsymbol{q}_N, \boldsymbol{q}_{N-1})\delta\boldsymbol{q}_N.
\end{aligned} \tag{4.1.20}$$

固定端点, 得到离散的 Euler-Lagrange 方程组

$$\frac{\partial}{\partial\boldsymbol{q}_1}L(\boldsymbol{q}_1, \boldsymbol{q}_0) + \frac{\partial}{\partial\boldsymbol{q}_0}L(\boldsymbol{q}_2, \boldsymbol{q}_1) = \boldsymbol{0}. \tag{4.1.21}$$

我们从离散的变分原理可以得到 $M \times M$ 上的 Lagrange 2-形式. 变分 (4.1.20) 中的两个边界项定义了 $M \times M$ 上两个离散的 Lagrange 1-形式:

$$\theta_L^-(\boldsymbol{q}_1, \boldsymbol{q}_0) \cdot (\delta\boldsymbol{q}_1, \delta\boldsymbol{q}_0) \triangleq \frac{\partial}{\partial\boldsymbol{q}_0}L(\boldsymbol{q}_1, \boldsymbol{q}_0)\delta\boldsymbol{q}_0, \tag{4.1.22}$$

$$\theta_L^+(\boldsymbol{q}_1, \boldsymbol{q}_0) \cdot (\delta\boldsymbol{q}_1, \delta\boldsymbol{q}_0) \triangleq \frac{\partial}{\partial\boldsymbol{q}_1}L(\boldsymbol{q}_1, \boldsymbol{q}_0)\delta\boldsymbol{q}_1. \tag{4.1.23}$$

θ_L^- 和 θ_L^+ 可视做 Lagrange 1-形式的离散形式. 把离散泛函变分限制在离散的变分解空间上, 假设离散解流为 F_{L_d}, 则变分 (4.1.20) 导出

$$\mathrm{d}\mathcal{L} = \theta_L^- + F_{L_d}^*\theta_L^+. \tag{4.1.24}$$

又因为

$$\theta_L^- + \theta_L^+ = \mathrm{d}L, \tag{4.1.25}$$

所以对其外微分有

$$\mathrm{d}\theta_L^- + \mathrm{d}\theta_L^+ = 0. \tag{4.1.26}$$

那么两个离散的 1-形式可以定义唯一的离散 Lagrange 2-形式 ω_L:

$$\omega_L = \mathrm{d}\theta_L^- = -\mathrm{d}\theta_L^+. \tag{4.1.27}$$

对 (4.1.24) 式作外微分, 得到辛流公式

$$F_{L_d}^* \omega_L = \omega_L. \tag{4.1.28}$$

上面简单回顾了不受外力作用的 Lagrange 系统的 Veselov 离散变分, 现在我们考虑受外力作用的 Lagrange 系统的离散变分. 定义两个离散的 Lagrange 力 $f_d^+, f_d^- : M \times M \to T^*M$, 它们是保纤维的, 即

$$\begin{aligned} f_d^+(\boldsymbol{q}_0, \boldsymbol{q}_1) &= (\boldsymbol{q}_1, f_d^+(\boldsymbol{q}_0, \boldsymbol{q}_1)), \\ f_d^-(\boldsymbol{q}_0, \boldsymbol{q}_1) &= (\boldsymbol{q}_1, f_{\boldsymbol{d}}^-(\boldsymbol{q}_0, \boldsymbol{q}_1)). \end{aligned} \tag{4.1.29}$$

将上面的两个离散力组合起来得到 $T^*(M \times M)$ 上的一个 1-形式 $f_d : M \times M \to T^*(M \times M)$:

$$f_d(\boldsymbol{q}_0, \boldsymbol{q}_1) \cdot (\delta \boldsymbol{q}_0, \delta \boldsymbol{q}_1) = f_d^+(\boldsymbol{q}_0, \boldsymbol{q}_1)\delta \boldsymbol{q}_1 + f_d^-(\boldsymbol{q}_0, \boldsymbol{q}_1)\delta \boldsymbol{q}_0. \tag{4.1.30}$$

和离散的 Lagrange 函数一样, 离散的 Lagrange 力也依赖于时间步长 τ. 定义了两个离散力之后, 得到受外力作用的离散 Lagrange-d'Alembert 原理: 寻找离散曲线 $\{\boldsymbol{q}_k\}_{k=0}^N$, 使得对任意固定端点的变分 $\{\delta \boldsymbol{q}_k\}_{k=0}^N$, 成立

$$\delta \sum_{k=0}^{N-1} L_d(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}) + \sum_{k=0}^{N-1} f_d(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}) \cdot (\delta \boldsymbol{q}_k, \delta \boldsymbol{q}_{k+1}) = 0, \tag{4.1.31}$$

其中 L_d 是 Lagrange 函数 L 的离散形式, 从而得到新的离散 Euler-Lagrange 方程

$$\mathrm{D}_2 L_d(\boldsymbol{q}_{k-1}, \boldsymbol{q}_k) + \mathrm{D}_1 L_d(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}) + f_d^+(\boldsymbol{q}_{k-1}, \boldsymbol{q}_k) + f_d^-(\boldsymbol{q}_k, \boldsymbol{q}_{k+1}) = 0, \tag{4.1.32}$$

这里 $\mathrm{D}_1 = \dfrac{\partial}{\partial \boldsymbol{q}_0}$, $\mathrm{D}_2 = \dfrac{\partial}{\partial \boldsymbol{q}_1}$. 方程 (4.1.32) 确定离散的 Lagrange 映射

$$F_{L_d} : M \times M \to M \times M.$$

4.1.4 受外力作用的离散 Legendre 变换和 Noether 定理

在连续的情况下, 处理受外力作用和不受外力作用的 Lagrange 系统均采用标准的 Legendre 变换; 而在离散的情况下, 这两类 Lagrange 系统需要采用不同的 Legendre 变换. 对于受外力作用的 Lagrange 系统, 在构造离散的 Legendre 变换时要考虑外力的影响. 定义两个离散 Legendre 变换:

$$\begin{aligned}
&\mathbb{F}^{f^+}L_d:(\boldsymbol{q}_0,\boldsymbol{q}_1)\mapsto(\boldsymbol{q}_1,\boldsymbol{p}_1)=(\boldsymbol{q}_1,\mathrm{D}_2L_d(\boldsymbol{q}_0,\boldsymbol{q}_1)+f_d^+(\boldsymbol{q}_0,\boldsymbol{q}_1)),\\
&\mathbb{F}^{f^-}L_d:(\boldsymbol{q}_0,\boldsymbol{q}_1)\mapsto(\boldsymbol{q}_0,\boldsymbol{p}_0)=(\boldsymbol{q}_0,-\mathrm{D}_1L_d(\boldsymbol{q}_0,\boldsymbol{q}_1)-f_d^-(\boldsymbol{q}_0,\boldsymbol{q}_1)).
\end{aligned}\tag{4.1.33}$$

根据上面的定义和离散 Euler-Lagrange 方程 (4.1.32), 对应的离散 Hamilton 系统存在 Hamilton 映射 $\widetilde{F}_{H_d}:(\boldsymbol{q}_0,\boldsymbol{p}_0)\mapsto(\boldsymbol{q}_1,\boldsymbol{p}_1), \widetilde{F}_{H_d}=\mathbb{F}^{f^\pm}L_d\circ F_{L_d}\circ(\mathbb{F}^{f^\pm}L_d)^{-1}$, 其中 $\boldsymbol{p}_0$ 和 $\boldsymbol{p}_1$ 由离散 Legendre 变换给出:

$$\boldsymbol{p}_0=-\mathrm{D}_1L_d(\boldsymbol{q}_0,\boldsymbol{q}_1)-f_d^-(\boldsymbol{q}_0,\boldsymbol{q}_1),\quad \boldsymbol{p}_1=\mathrm{D}_2L_d(\boldsymbol{q}_0,\boldsymbol{q}_1)+f_d^+(\boldsymbol{q}_0,\boldsymbol{q}_1).\tag{4.1.34}$$

这是标准的离散 Hamilton 映射加上离散的外力.

和连续时的情况一样, 在群作用 $\boldsymbol{\Phi}:G\times M\to M$ 下, 假设离散的 Lagrange 函数 L_d 是不变的, 令变分 $\delta\boldsymbol{q}_k=\boldsymbol{\xi}_M(\boldsymbol{q}_k)$, 计算变分 (4.1.31), 得到

$$\begin{aligned}
&\sum_{k=0}^{N-1}\mathrm{d}L_d(\boldsymbol{q}_k,\boldsymbol{q}_{k+1})\cdot\boldsymbol{\xi}_{M\times M}(\boldsymbol{q}_k,\boldsymbol{q}_{k+1})\\
&\qquad+\sum_{k=0}^{N-1}f_d(\boldsymbol{q}_k,\boldsymbol{q}_{k+1})\cdot\boldsymbol{\xi}_{M\times M}(\boldsymbol{q}_k,\boldsymbol{q}_{k+1})\\
&=\sum_{k=0}^{N-1}f_d(\boldsymbol{q}_k,\boldsymbol{q}_{k+1})\cdot\boldsymbol{\xi}_{M\times M}(\boldsymbol{q}_k,\boldsymbol{q}_{k+1}),
\end{aligned}\tag{4.1.35}$$

而利用离散的分部积分直接计算变分 (4.1.31), 得到另外一种表达:

$$\sum_{k=1}^{N-1}[\mathrm{D}_2L_d(\boldsymbol{q}_{k-1},\boldsymbol{q}_k)+\mathrm{D}_1L_d(\boldsymbol{q}_k,\boldsymbol{q}_{k+1})$$
$$+f_d^+(\boldsymbol{q}_{k-1},\boldsymbol{q}_k)+f_d^-(\boldsymbol{q}_k,\boldsymbol{q}_{k+1})]\cdot\boldsymbol{\xi}_M(\boldsymbol{q}_k)$$
$$+[\mathrm{D}_2L_d(\boldsymbol{q}_{N-1},\boldsymbol{q}_N)+f_d^+(\boldsymbol{q}_{N-1},\boldsymbol{q}_N)]\cdot\boldsymbol{\xi}_M(\boldsymbol{q}_N)$$
$$+[\mathrm{D}_1L_d(\boldsymbol{q}_0,\boldsymbol{q}_1)+f_d^-(\boldsymbol{q}_0,\boldsymbol{q}_1)]\cdot\boldsymbol{\xi}_M(\boldsymbol{q}_0)$$
$$=\mathbb{F}^{f^+}L_d(\boldsymbol{q}_{N-1},\boldsymbol{q}_N)\cdot\boldsymbol{\xi}_M(\boldsymbol{q}_N)-\mathbb{F}^{f^-}L_d(\boldsymbol{q}_0,\boldsymbol{q}_1)\cdot\boldsymbol{\xi}_M(\boldsymbol{q}_0).\quad(4.1.36)$$

如何定义离散的 Lagrange 动量映射, 存在两种选择: 第一种, 根据连续时的定义 (4.1.14) 式, 利用离散的 Lagrange 1-形式 θ_L^+ 和 θ_L^- 来定义; 第二种, 利用离散的 Legendre 变换如下定义:

$$\begin{aligned}J_{L_d}^{f^+}(\boldsymbol{q}_0,\boldsymbol{q}_1)\cdot\boldsymbol{\xi}&=\langle\mathbb{F}^{f^+}L_d(\boldsymbol{q}_0,\boldsymbol{q}_1),\boldsymbol{\xi}_M(\boldsymbol{q}_1)\rangle,\\J_{L_d}^{f^-}(\boldsymbol{q}_0,\boldsymbol{q}_1)\cdot\boldsymbol{\xi}&=\langle\mathbb{F}^{f^-}L_d(\boldsymbol{q}_0,\boldsymbol{q}_1),\boldsymbol{\xi}_M(\boldsymbol{q}_1)\rangle.\end{aligned}\quad(4.1.37)$$

对于受外力作用的连续系统和不受外力作用的系统 (包括连续和离散系统), 利用 Legendre 变换定义的 Lagrange 动量映射和利用 Lagrange 1-形式定义的动量映射是等价的. 然而, 对于受外力作用的离散系统, 两种定义是不同的. 在后面会说明受外力作用的离散 Lagrange 动量映射正确的定义为 (4.1.37).

取定义 (4.1.37), 要求变分 (4.1.35) 等于变分 (4.1.36), 则有

$$\left(J_{L_d}^{f^+}\circ F_{L_d}^{N-1}-J_{L_d}^{f-}\right)(\boldsymbol{q}_0,\boldsymbol{q}_1)\cdot\boldsymbol{\xi}$$
$$=\sum_{k=0}^{N-1}f_d(\boldsymbol{q}_k,\boldsymbol{q}_{k+1})\cdot\boldsymbol{\xi}_{M\times M}(\boldsymbol{q}_k,\boldsymbol{q}_{k+1}).\quad(4.1.38)$$

上式描述了离散 Lagrange 动量映射的发展变化. 如果离散力垂直于群作用, 那么对 $\forall\boldsymbol{\xi}\in\mathcal{G}$, 成立

$$0=\langle\mathrm{d}L_d+f_d,\boldsymbol{\xi}_{M\times M}\rangle=J_{L_d}^{f^+}-J_{L_d}^{f^-},\quad(4.1.39)$$

即两个离散的 Lagrange 动量映射是相等的, 记为 $J_{L_d}^f:M\times M\to\mathcal{G}^*$. 由上面的推导得到如下受外力作用的离散情形的 Noether 定理:

定理 4.1.2 设离散的 Lagrange 系统 $L_d : M\times M \to \mathbf{R}$ 受到离散外力 $f_d^+, f_d^- : M\times M \to T^*M$ 作用, 又设存在群作用 $\boldsymbol{\Phi} : G\times M \to M$, 使得对 $\forall \boldsymbol{\xi} \in \mathcal{G}$, 有

$$\langle f_d, \boldsymbol{\xi}_{M\times M}\rangle = 0,$$

则 Lagrange 离散解流 F_{L_d} 保持离散的 Lagrange 动量映射 $J_{L_d}^f : M\times M \to \mathcal{G}^*$, 即

$$J_{L_d}^f \circ F_{L_d} = J_{L_d}^f.$$

Lagrange 动量映射是 Hamilton 动量映射在 Legendre 变换下的拉回, 受外力作用的离散系统也是如此, 所以 Noether 定理对受外力作用的离散 Hamilton 系统也成立, 此时离散 Hamilton 相流 $\widetilde{F}_{H_d}$ 保持离散的典则动量映射 $J_H : T^*M \to \mathcal{G}^*$. 和连续时一样, 它们也不保持离散辛形式.

以上介绍的都是受外力作用的 Lagrange 系统的离散力学. 约束系统的离散力学和变分积分子在文献 [MW01], [Arn63], [Mos62] 中都有所研究. 关于受外力作用的系统的研究, 像本章所介绍的, 基本上都是一个二阶 Lagrange 泛函变分附加一个外力变分, 或者标准的 Hamilton 系统的动量方程附加一个外力项, 再通过 Legendre 变换将二者联系起来.

在后面的章节中, 我们把 Hamilton 系统推广到受外力作用的系统, 即文献 [San83] 中介绍的 Birkhoff 系统, 进而推广到 Birkhoff 偏微分系统, 研究它们的离散力学和场论以及辛格式和变分积分子的构造, 并研究系统的几何结构和外力作用之间的关系. 从方程结构来看, Birkhoff 系统比以上介绍的 Hamilton 系统更普遍, 而且在牛顿力学、空间力学、统计力学、工程及生物物理等学科中应用广泛. 有三点可以说明 Birkhoff 系统是以上 Hamilton 系统的自然推广:

(1) 允许一个更一般的一阶作用泛函, 即 Pfaff 作用泛函;

(2) 具有一个更一般的 Lie 代数结构;

(3) 在相空间和包含了底空间的增广空间切空间上分别存在各自的几何结构, 这两种结构是由外力联系起来的, 并且相流和增广解流分别保持这两种几何结构.

所以研究 Birkhoff 系统的几何结构和数值方法有更加广泛的意义.

§4.2 Pfaff 作用泛函的几何变分

我们考虑一个 n 维微分流形 Z, 它的切丛是 TZ. 设 Z 和 TZ 的局部坐标分别为 z^ν 和 $(z^\nu, \dot{z}^\nu)$, $\nu = 1, \cdots, n$, 并假定 n 为偶数. 通过一类函数 $L: TZ \to \mathbf{R}$ 可以构造一类作用在 Z 的 C^1 曲线 $\boldsymbol{z}(t)$ 上的作用泛函 $\mathcal{L}$:

$$\mathcal{L}(\boldsymbol{z}(t)) = \int_a^b L(\boldsymbol{z}(t), \dot{\boldsymbol{z}}(t))\mathrm{d}t. \tag{4.2.1}$$

如果函数 L 可以写成如下形式:

$$L = F_\nu(\boldsymbol{z}, t)\dot{z}^\nu - B(\boldsymbol{z}, t), \tag{4.2.2}$$

其中 $\boldsymbol{F}(\boldsymbol{z}, t) = (F_1(\boldsymbol{z}, t), \cdots, F_n(\boldsymbol{z}, t))^{\mathrm{T}}$ 是向量函数, $B(\boldsymbol{z}, t)$ 是标量函数, 则泛函 $\mathcal{L}(\boldsymbol{z}(t))$ 就是所谓的 **Pfaff 作用泛函**[San83], 它所对应的积分函数 L 是处处奇异的, 即 $\dfrac{\partial^2 L}{\partial \dot{z}^\mu \dot{z}^\nu} = 0$. Pfaff 作用泛函问题对应的变分微分方程就是一类 Birkhoff 方程组. 在本书后面几章中, 我们将着重从微分方程出发讨论它们的辛结构和多辛结构, 并且利用生成函数和生成泛函构造辛积分子. 本章我们讨论其变分积分子.

首先回忆一下在讨论变分微分方程的辛结构时经典的几何处理方法. Lagrange 函数 L 本身定义了一个从流形 M 的切丛 TM 到余切丛 T^*M 的丛映射 $\mathbb{F}L: TM \to T^*M$, 即 Legendre 变换:

$$\mathbb{F}L(\boldsymbol{v_q})\boldsymbol{w_q} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(\boldsymbol{v_q} + \varepsilon \boldsymbol{w_q})\Big|_{\varepsilon=0}, \quad \boldsymbol{v_q}, \boldsymbol{w_q} \in TM. \tag{4.2.3}$$

而通过典则的投影映射可以很自然地在 T^*M 上定义典则 1-形式 θ, 再通过 $\mathbb{F}L$ 诱导的拉回映射, 将 θ 拉回到 TM 上, 从而得到 TM 上的 Lagrange 1-形式 θ_L, 进而得到 Lagrange 2-形式 $\omega_L = -\mathrm{d}\theta_L$.

然而, 用以上的方法处理 Pfaff 作用泛函却行不通, 因为其积分函数 L 是处处奇异的, 无法定义类似上面的丛映射 $\mathbb{F}L$. 因此, 下面我们不采用经典的拉回方法, 而是直接通过变分运算来分析系统结构.

定义 $T(TZ)$ 的一个二阶子流形:

$$\ddot{Z} \triangleq \{\boldsymbol{w} \in T(TZ) | T_{\pi_Z}(\boldsymbol{w}) = \pi_{TZ}(\boldsymbol{w}) \subset T(TZ)\},$$

这里 $\pi_{TZ} : T(TZ) \to TZ$ 是投影映射 $\pi_Z : TZ \to Z$ 在 TZ 上的提升映射. $T(TZ)$ 的局部坐标为 $((\boldsymbol{z}, \dot{\boldsymbol{z}}), (\dot{\boldsymbol{z}}, \ddot{\boldsymbol{z}}))$. 对于非奇异的 Lagrange 函数 L, 有以下结论成立:

定理 4.2.1 假设 L 是 C^k $(k \geqslant 2)$ 的, 那么存在唯一的 C^{k-2} 映射 $D_{EL}L : \ddot{Z} \to T^*Z$ 和 TZ 上唯一的 C^{k-1} 1-形式 θ_L, 使得对所有的变分 $\boldsymbol{z}_\varepsilon(t) \in C^2(Z)$, 有

$$\mathrm{d}\pounds(\boldsymbol{z}(t))\delta\boldsymbol{z}(t) = \int_a^b D_{EL}L\left((\boldsymbol{z}, \dot{\boldsymbol{z}}), (\dot{\boldsymbol{z}}, \ddot{\boldsymbol{z}})\right)\delta\boldsymbol{z}\mathrm{d}t + \theta_L(\boldsymbol{z}, \dot{\boldsymbol{z}})\widehat{\delta}\boldsymbol{z}\Big|_a^b,$$
$$\delta\boldsymbol{z} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\boldsymbol{z}_\varepsilon(t)\Big|_{\varepsilon=0}, \quad \delta\dot{\boldsymbol{z}} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\dot{\boldsymbol{z}}_\varepsilon(t)\Big|_{\varepsilon=0}, \quad \widehat{\delta}\boldsymbol{z} = (\delta\boldsymbol{z}, \delta\dot{\boldsymbol{z}}), \tag{4.2.4}$$

其中 D_{EL} 为 Euler-Lagrange 映射, θ_L 为 Lagrange 1-形式.

这个结论及其证明可参见文献 [MW01], 我们稍做了修改. 这里的 Euler-Lagrange 映射即 Euler 算子作用到 Lagrange 函数, $D_{EL}L((\boldsymbol{z}, \dot{\boldsymbol{z}}), (\dot{\boldsymbol{z}}, \ddot{\boldsymbol{z}})) = 0$ 即 Euler-Lagrange 方程. 下面我们就利用这种几何变分的思想处理 Pfaff 变分问题, 得到 TZ 子空间上的 Birkhoff 型的 Euler-Lagrange 一阶方程. 就局部坐标直接对 Pfaff 作用泛函 $\pounds(\boldsymbol{z}(t))$ 变分, 得到

$$\begin{aligned}\mathrm{d}\pounds(\boldsymbol{z}(t))\delta\boldsymbol{z}(t) &= \int_a^b \delta z^\nu \left(\frac{\partial L}{\partial z^\nu} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{z}^\nu}\right)\mathrm{d}t + \frac{\partial L}{\partial \dot{z}^\nu}\delta z^\nu\Big|_a^b \\ &= \int_a^b \delta z^\mu \left[\left(\frac{\partial F_\nu}{\partial z^\mu} - \frac{\partial F_\mu}{\partial z^\nu}\right)\dot{z}^\nu - \frac{\partial B}{\partial z^\mu} - \frac{\partial F_\mu}{\partial t}\right]\mathrm{d}t + F_\mu \delta z^\mu\Big|_a^b.\end{aligned} \tag{4.2.5}$$

有上面的局部坐标上的变分做基础, 我们导出了定理 4.2.1 的退化形式:

定理 4.2.2 假设 $\boldsymbol{F}(\boldsymbol{z}, t)$ 和 $B(\boldsymbol{z}, t)$ 是 C^k $(k \geqslant 1)$ 的, 那么存在唯一的 C^{k-1} 映射 $D_{EL}L : TZ \to T^*Z$, 并且在 TZ 上存在唯一的 C^k

1-形式 θ_L, 使得对所有的变分 $\boldsymbol{z}_\varepsilon(t)\in C^2(Z)$, 有

$$\mathrm{d}\mathcal{L}(\boldsymbol{z}(t))\delta\boldsymbol{z}(t)=\int_a^b D_{EL}L(\boldsymbol{z},\dot{\boldsymbol{z}})\delta\boldsymbol{z}\mathrm{d}t+\theta_L(\boldsymbol{z})\widehat{\delta}\boldsymbol{z}\Big|_a^b, \tag{4.2.6}$$

其中 $\theta_L(\boldsymbol{z})\widehat{\delta}\boldsymbol{z}$ 在局部坐标意义下表示为 $F_\mu\delta z^\mu$.

我们沿用 Lagrange 1-形式的叫法, 称此处的 θ_L 为 **Pfaff 1-形式**.

由以上变分得到的 Euler-Lagrange 方程是 Birkhoff 型变分微分方程. Pfaff 作用泛函不依赖于局部坐标, 所以由上面局部坐标上的变分结果, 全局的存在性和唯一性得证.

将由以上变分得到的 Birkhoff 型 Euler-Lagrange 方程对应的相流记为 F_t. 我们设 $\mathcal{C}_L$ 是变分问题解的一个子空间, 对应初始条件 $\boldsymbol{v}_{\boldsymbol{z}}=(\boldsymbol{z}(0),\dot{\boldsymbol{z}}(0))$. 将 $\mathcal{L}$ 限制在 $\mathcal{C}_L$ 上, 这时 $\mathcal{C}_L$ 上存在积分曲线 $\xi\mapsto F_\xi(\boldsymbol{v}_{\boldsymbol{z}})=(\boldsymbol{z}(\xi),\dot{\boldsymbol{z}}(\xi))$, $\xi\in[0,t]$. 在 $[0,t]$ 上定义泛函

$$\mathcal{L}_t(\boldsymbol{v}_{\boldsymbol{z}})=\int_0^t L(\boldsymbol{z}(\xi),\dot{\boldsymbol{z}}(\xi))\mathrm{d}\xi. \tag{4.2.7}$$

对 $\mathcal{L}_t(\boldsymbol{v}_{\boldsymbol{z}})$ 变分, Euler-Lagrange 方程成立, 所以仅剩下两个边界项:

$$\mathrm{d}\mathcal{L}_t(\boldsymbol{v}_{\boldsymbol{z}})\boldsymbol{w}_{\boldsymbol{v}_{\boldsymbol{z}}}=\theta_L(F_t(\boldsymbol{v}_{\boldsymbol{z}}))\cdot\frac{\mathrm{d}}{\mathrm{d}\varepsilon}F_t(\boldsymbol{v}_{\boldsymbol{z}}^\varepsilon)\Big|_{\varepsilon=0}-\theta_L(\boldsymbol{v}_{\boldsymbol{z}})\cdot\boldsymbol{w}_{\boldsymbol{v}_{\boldsymbol{z}}}, \tag{4.2.8}$$

其中 TZ 上的曲线 $\varepsilon\mapsto\boldsymbol{v}_{\boldsymbol{z}}^\varepsilon$ 满足 $\boldsymbol{v}_{\boldsymbol{z}}^0=\boldsymbol{v}_{\boldsymbol{z}}$, 且 $\boldsymbol{w}_{\boldsymbol{v}_{\boldsymbol{z}}}=\dfrac{\mathrm{d}}{\mathrm{d}\varepsilon}\boldsymbol{v}_{\boldsymbol{z}}^\varepsilon\Big|_{\varepsilon=0}$ 是其变分. 因此得到等式

$$\mathrm{d}\mathcal{L}_t=F_t^*\theta_L-\theta_L, \tag{4.2.9}$$

从而得到辛流公式

$$0=F_t^*(\mathrm{d}\theta_L)-\mathrm{d}\theta_L. \tag{4.2.10}$$

上述过程得到的 Pfaff 1-形式和辛流公式只能说描述了 Pfaff 变分问题的部分结构. 事实上, 几何变分方法对于具体的问题结果不同的一个原因是考虑的几何空间不同, 但从局部坐标来看, 变分计算的过程是一样的. 下面我们就 Pfaff 作用泛函再讨论和比较一下另外两种不同的变分. 由于 Pfaff 变分问题所对应的 Birkhoff 型 Euler-Lagrange 方程带有外力项, 而外力项可以看做由于系统的结构依赖于自变量而

产生的, 所以后面两种变分方法因为考虑了自变量, 更适合 Birkhoff 型 Euler-Lagrange 方程.

首先对上面的几何空间加以调整. 原来的相空间增加一维, 扩展为 $\widehat{Z}$, 其局部坐标为 $(\widehat{z}^0, \widehat{z}^1, \cdots, \widehat{z}^n)$, 其中 $\widehat{z}^0 = t \triangleq s(t)$, $\widehat{z}^\nu = z^\nu(\nu = 1, \cdots, n)$. 它的切空间 $T\widehat{Z}$ 的局部坐标为 $(\widehat{z}^\nu, \dot{\widehat{z}}^\nu), \nu = 0, 1, \cdots, n$. 变分问题的目标空间现在就变成了 $T\widehat{Z}$, 原来的作用泛函改写为

$$\mathcal{L}(\widehat{\boldsymbol{z}}(t)) = \int_a^b [F_\nu(\boldsymbol{z}, s)\dot{z}^\nu - B(\boldsymbol{z}, s)\dot{s}]\,\mathrm{d}t. \tag{4.2.11}$$

这个作用泛函依旧是奇异的, 但它同样可以变分得到一组微分方程组, 而且这样的方程组和原来的 Birkhoff 型 Euler-Lagrange 方程是等价的. 对上面的作用泛函关于增广空间 $\widehat{Z}$ 变分, 得到

$$\begin{aligned}\widehat{\mathrm{d}}\mathcal{L}(\widehat{\boldsymbol{z}}(t))\delta\widehat{\boldsymbol{z}} = \int_a^b \Bigg[&\delta z^\mu \left(\frac{\partial}{\partial z^\mu}F_\nu(\boldsymbol{z}, s)\dot{z}^\nu - \frac{\partial}{\partial z^\mu}B(\boldsymbol{z}, s)\dot{s}\right) \\ &+\delta s \left(\frac{\partial}{\partial s}F_\nu(\boldsymbol{z}, s)\dot{z}^\nu - \frac{\partial}{\partial s}B(\boldsymbol{z}, s)\dot{s}\right) \\ &+F_\mu(\boldsymbol{z}, s)\delta\dot{z}^\mu - B(\boldsymbol{z}, s)\delta\dot{s}\Bigg]\mathrm{d}t.\end{aligned} \tag{4.2.12}$$

上式最后两项利用分部积分法, 得

$$\begin{aligned}&\int_a^b (F_\mu(\boldsymbol{z}, s)\delta\dot{z}^\mu - B(\boldsymbol{z}, s)\delta\dot{s})\,\mathrm{d}t \\ &= (F_\mu\delta z^\mu - B\delta s)\,\big|_a^b - \int_a^b \left(\frac{\mathrm{d}F_\mu}{\mathrm{d}t}\delta z^\mu - \frac{\mathrm{d}B}{\mathrm{d}t}\delta s\right)\mathrm{d}t,\end{aligned} \tag{4.2.13}$$

又

$$\frac{\mathrm{d}F_\mu}{\mathrm{d}t} = \frac{\partial F_\mu}{\partial z^\nu}\dot{z}^\nu + \frac{\partial F_\mu}{\partial s}\dot{s}, \quad \frac{\mathrm{d}B}{\mathrm{d}t} = \frac{\partial B}{\partial z^\mu}\dot{z}^\mu + \frac{\partial B}{\partial s}\dot{s}, \tag{4.2.14}$$

所以

$$\begin{aligned}\widehat{\mathrm{d}}\mathcal{L}(\widehat{\boldsymbol{z}}(t))\delta\widehat{\boldsymbol{z}} = &[F_\mu\delta z^\mu - B\delta s]\,\big|_a^b \\ &+\int_a^b \left\{\delta z^\mu \left[\left(\frac{\partial F_\nu}{\partial z^\mu} - \frac{\partial F_\mu}{\partial z^\nu}\right)\dot{z}^\nu - \left(\frac{\partial B}{\partial z^\mu} + \frac{\partial F_\mu}{\partial s}\right)\dot{s}\right]\right. \\ &\left.+\delta s\left[\frac{\partial F_\mu}{\partial s}\dot{z}^\mu + \frac{\partial B}{\partial z^\mu}\dot{z}^\mu\right]\right\}\mathrm{d}t.\end{aligned} \tag{4.2.15}$$

定理 4.2.2 对以上扩展依旧适用. 注意上面的变分过程中的外微分记号 d 表示针对 Z 变分, 而 $\widehat{\mathrm{d}}$ 表示在增广空间 $\widehat{Z}$ 上变分, 不同于直接在相空间 Z 上的变分. 设扩展的变分问题存在变分解流 $\xi \mapsto \widehat{F}_\xi(\boldsymbol{v}_{\widehat{\boldsymbol{z}}})$, $\xi \in [0,t]$, $\boldsymbol{v}_{\widehat{\boldsymbol{z}}} \in T\widehat{Z}$. 将作用泛函 $\mathcal{L}$ 定义在这一段积分曲线上, 得到 $\mathcal{L}_t : T\widehat{Z} \to \mathbf{R}$:

$$\mathcal{L}_t(\boldsymbol{v}_{\widehat{\boldsymbol{z}}}) = \int_0^t L(\widehat{\boldsymbol{z}}(\xi), \dot{\widehat{\boldsymbol{z}}}(\xi))\mathrm{d}\xi. \tag{4.2.16}$$

对此作用泛函作几何变分, 得到

$$\widehat{\mathrm{d}}\mathcal{L}_t(\boldsymbol{v}_{\widehat{\boldsymbol{z}}})\boldsymbol{w}_{\boldsymbol{v}_{\widehat{\boldsymbol{z}}}} = \widehat{\theta}_L(\widehat{F}_t(\boldsymbol{v}_{\widehat{\boldsymbol{z}}}))\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\widehat{F}_t(\boldsymbol{v}^\varepsilon_{\widehat{\boldsymbol{z}}})\right|_{\varepsilon=0} - \widehat{\theta}_L(\boldsymbol{v}_{\widehat{\boldsymbol{z}}})\boldsymbol{w}_{\boldsymbol{v}_{\widehat{\boldsymbol{z}}}}, \tag{4.2.17}$$

其中 $\varepsilon \mapsto \boldsymbol{v}^\varepsilon_{\widehat{\boldsymbol{z}}}$ 是 $T\widehat{Z}$ 上的任意曲线, 满足 $\boldsymbol{v}^0_{\widehat{\boldsymbol{z}}} = \boldsymbol{v}_{\widehat{\boldsymbol{z}}}, \left.\dfrac{\mathrm{d}}{\mathrm{d}\varepsilon}\boldsymbol{v}^\varepsilon_{\widehat{\boldsymbol{z}}}\right|_{\varepsilon=0} = \boldsymbol{w}_{\boldsymbol{v}_{\widehat{\boldsymbol{z}}}}$, 所以有

$$\widehat{\mathrm{d}}\mathcal{L}_t = \widehat{F}_t^*\widehat{\theta}_L - \widehat{\theta}_L. \tag{4.2.18}$$

对上式作用外微分 $\widehat{\mathrm{d}}$, 导出

$$\widehat{F}_t^*\widehat{\omega}_L = \widehat{\omega}_L, \quad \widehat{\omega}_L = -\widehat{\mathrm{d}}\widehat{\theta}_L. \tag{4.2.19}$$

具体到局部坐标上, 将变分 (4.2.15) 固定端点, 得到带有外力项 $\dfrac{\partial \boldsymbol{F}}{\partial t}$ 的 Birkhoff 型 Euler-Lagrange 方程

$$\left(\frac{\partial F_\nu}{\partial z^\mu} - \frac{\partial F_\mu}{\partial z^\nu}\right)\dot{z}^\nu - \left(\frac{\partial B}{\partial z^\mu} + \frac{\partial F_\mu}{\partial s}\right)\dot{s} = 0. \tag{4.2.20}$$

将 $s = t$ 代入, 它等价于第一种变分得到的 Euler-Lagrange 方程. 在 δs 方向得到一个方程:

$$\frac{\partial F_\mu}{\partial s}\dot{z}^\mu + \frac{\partial B}{\partial z^\mu}\dot{z}^\mu = 0. \tag{4.2.21}$$

假设 Euler-Lagrange 方程 (4.2.20) 中一阶项的系数矩阵 $\left(\dfrac{\partial F_\nu}{\partial z^\mu} - \dfrac{\partial F_\mu}{\partial z^\nu}\right)$ 是非奇异的, 则由于它反对称, 方程 (4.2.20) 左乘 $\dot{z}^\mu$ 就可得到方程 (4.2.21). 所以方程 (4.2.21) 是自然成立的. 我们称其为**形式能量公式**. 以边界项定义 Pfaff 1-形式为

$$\widehat{\theta}_L = F_\mu(\boldsymbol{z},t)\widehat{\mathrm{d}}z^\mu - B(\boldsymbol{z},s)\widehat{\mathrm{d}}s, \tag{4.2.22}$$

定义 Pfaff 2-形式为 $\widehat{\omega}_L = -\widehat{\mathrm{d}}\widehat{\theta}_L$, 即

$$\widehat{\omega}_L = \frac{\partial F_\nu}{\partial z^\mu}\mathrm{d}z^\mu \wedge \mathrm{d}z^\nu + \frac{\partial F_\mu}{\partial s}\mathrm{d}s \wedge \mathrm{d}z^\mu + \frac{\partial B}{\partial z^\mu}\mathrm{d}s \wedge \mathrm{d}z^\mu. \tag{4.2.23}$$

因为增广空间 $\widehat{Z}$ 存在解流 $\widehat{F}_t$, 所以可以导出

$$\widehat{F}_t^* \widehat{\omega}_L = \widehat{\omega}_L. \tag{4.2.24}$$

令

$$\theta_L = F_\mu(\boldsymbol{z}, t)\mathrm{d}z^\mu, \quad \omega_L = -\mathrm{d}\theta_L, \tag{4.2.25}$$

那么 $\widehat{\omega}_L = \omega_L + E_\mu \mathrm{d}s \wedge \mathrm{d}z^\mu$, $E_\mu \mathrm{d}s \wedge \mathrm{d}z^\mu = \dfrac{\partial F_\mu}{\partial s}\mathrm{d}s \wedge \mathrm{d}z^\mu + \dfrac{\partial B}{\partial z^\mu}\mathrm{d}s \wedge \mathrm{d}z^\mu$, 则

$$\omega_L + E_\mu \mathrm{d}s \wedge \mathrm{d}z_\mu = \widehat{\omega}_L(\boldsymbol{z}, s) = \widehat{F}_t^* \widehat{\omega}_L(\boldsymbol{z}, s) = \widehat{F}_t^* \omega_L + \widehat{F}_t^*(E_\mu \mathrm{d}s \wedge \mathrm{d}z^\mu). \tag{4.2.26}$$

根据约束条件 (4.2.21), 知道 $\widehat{F}_t^*(E_\mu \mathrm{d}s \wedge \mathrm{d}z^\mu) = E_\mu \mathrm{d}(s+\xi) \wedge \mathrm{d}z^\mu$, 所以

$$\widehat{F}_t^* \omega_L = \omega_L - E_\mu \mathrm{d}\xi \wedge \mathrm{d}z^\mu. \tag{4.2.27}$$

因为 $s = t$, 所以在相空间 Z 上存在相流 F_t , F_t 包含于 $\widehat{F}_t$. 因此, 我们可以假设存在一个 2-形式 ω_s , 它形如某个 1-形式外乘 $\mathrm{d}s$, 使得

$$\widehat{F}_t^* \omega_L = F_t^* \omega_L + \omega_s. \tag{4.2.28}$$

对比 (4.2.27) 式和 (4.2.28) 式, 得到相空间的**约束辛流公式**

$$F_t^* \omega_L = \omega_L, \quad \omega_s = -E_\mu \mathrm{d}\xi \wedge \mathrm{d}z^\mu \ (\text{约束条件}). \tag{4.2.29}$$

如果直接在相空间 Z 上讨论变分, 也可以得到相空间的辛流公式 $F_t^* \omega_L = \omega_L$, 但是不能得到形式能量公式. 在第五章中, 我们将从方程对称性的角度讨论, 给出具体的约束条件.

辛结构 ω_L 包含于 $\widehat{\omega}_L$, 我们称前者为相空间辛结构, 后者可以看做全空间或者增广空间的广义辛结构. 在第五章和第六章通过对方程的对称性讨论, 这一点会得到更简洁的说明.

Birkhoff 系统的相空间辛结构必定显式依赖于时间参数 t, 而 Hamilton 系统无论是自治的还是非自治的, 其相空间辛结构一般不会显式依赖于时间和空间变量, 但是也存在很特别的 Hamilton 辛结构是显式依赖于时间和空间变量的. 这一点在后面两章中会有更加具体的说明. 虽然 Birkhoff 系统和 Hamilton 系统最大的区别在于, 前者因为辛结构 $\boldsymbol{F}$ 和能量函数 B 显式依赖于时间 t 而分别导致外力项 $\dfrac{\partial \boldsymbol{F}}{\partial t}$ 和能量不守恒, 而非自治 Hamilton 系统只存在能量不守恒, 其结构的耗散不存在或者是内部抵消了, 但是也可以说 Birkhoff 系统和不守恒的 Hamilton 系统的共同点也正在于此: 都是因为显式依赖于时间 t, 导致了结构或者 (和) 能量的不守恒.

下面我们利用文献[CGW03] 中提出的全变分方法对 Pfaff 作用泛函进行全变分, 与上面的结果做一个比较. 在增广空间 $\widehat{Z}$ 上变分的好处是, 虽然增加了一个变量 s, 但变分仍然类似在相空间上进行, 不涉及积分区域的变分. 而全变分会涉及积分区域的变分. 因此, 和全变分相比, 增广空间上的变分计算要简单些.

对 $\mathcal{L}$ 作全变分, 得到

$$
\begin{aligned}
&\mathrm{d}\mathcal{L}(\boldsymbol{z},t)\cdot(\delta\boldsymbol{z},\delta t)\\
&=\int_a^b\left[\left(\frac{\partial L}{\partial t}+\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot z^\nu}\dot z^\nu-L\right)\right)\delta t+\left(\frac{\partial L}{\partial z^\mu}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot z^\mu}\right)\right)\delta z^\mu\right]\mathrm{d}t\\
&\quad+\left[\left(L-\frac{\partial L}{\partial \dot z^\mu}\dot z^\mu\right)\delta t+\frac{\partial L}{\partial \dot z^\mu}\delta^z\mu\right]\Bigg|_a^b\\
&=\int_a^b\left[\left(\frac{\partial F_\nu}{\partial t}\dot z^\nu-\frac{\partial B}{\partial t}+\frac{\mathrm{d}B}{\mathrm{d}t}\right)\delta t+\left(\frac{\partial F_\nu}{\partial z^\mu}\dot z^\nu-\frac{\partial B}{\partial z^\mu}-\frac{\mathrm{d}}{\mathrm{d}t}F_\mu\right)\delta z^\mu\right]\mathrm{d}t\\
&\quad+(-B\delta t+F_\mu\delta z^\mu)\big|_a^b\\
&=\int_a^b\left[\left(\frac{\partial F_\nu}{\partial t}\dot z^\nu+\frac{\partial B}{\partial z^\nu}\dot z^\nu\right)\delta t+\left(\left(\frac{\partial F_\nu}{\partial z^\mu}-\frac{\partial F_\mu}{\partial z^\nu}\right)\dot z^\nu\right.\right.\\
&\quad\left.\left.-\frac{\partial B}{\partial z^\mu}-\frac{\partial F_\mu}{\partial t}\right)\delta z^\mu\right]\mathrm{d}t+(-B\delta t+F_\mu\delta z^\mu)\big|_a^b.
\end{aligned}
\tag{4.2.30}
$$

固定端点, 则沿 δz^μ 和 δt 方向得到带有外力项的 Birkhoff 型 Euler-Lagrange 方程 (4.2.20) 和形式能量公式 (4.2.21), 与增广空间上的变分

结果一样, 边界项也一样. 若令

$$\theta_L = -B\delta t + F_\mu \delta z^\mu,$$

我们同样可以得到在包含时间 t 的空间上的一个结构 $\omega_L = -\mathrm{d}\theta_L$, 使得包含时间 t 流动的相流 $\widehat{F}_t$ 保持这个结构. 若不考虑时间 t 的流动, 那么根据结构 ω_L 和约束公式, 我们也可以推出等价于 (4.2.29) 式的标准相流公式和相流约束条件. 这个结论和文献 [CGW03] 的结论有出入, 与 Marsden 等人的受外力作用的 Lagrange 流和 Hamilton 相流不保辛结构的结论也完全不同. 因为他们考虑的是增广解流是否保持相空间结构, 或者相流是否保持增广空间结构, 这两者当然都不能成立. 我们的结论是: 存在增广空间和标准相空间两个结构; 如果存在增广空间解流, 那么也存在标准相空间流, 后者包含于前者, 并且由增广解流保持增广空间结构可以推出标准相流保持标准相空间 (辛) 结构和一个相流约束条件 (把它们合称为约束辛流公式).

§4.3　Veselov 离散变分形式

我们再次回顾一下经典的 Veselov 离散[Ves88] 和 Marsden 变分形式的 Veselov 离散[Ves91].

用 $M \times M$ 作为相空间 M 的切丛 TM 的离散空间. 假定时间步长为 τ, 则 $M \times M$ 中的某一点 $(\boldsymbol{q}_1, \boldsymbol{q}_0)$ 对应切向量 $(\boldsymbol{q}_1 - \boldsymbol{q}_0)/\tau$. 定义一个光滑映射

$$L : M \times M \to \mathbf{R},$$
$$(\boldsymbol{q}_k, \boldsymbol{q}_{k-1}) \mapsto L(\boldsymbol{q}_k, \boldsymbol{q}_{k-1}),$$

离散的作用泛函定义为

$$\pounds = \sum_{k=1}^{n} L(\boldsymbol{q}_k, \boldsymbol{q}_{k-1}). \tag{4.3.1}$$

离散变分原理就是求 $(\boldsymbol{q}_0, \cdots, \boldsymbol{q}_n)$, 使得 $\pounds$ 取极值. 对上面的离散泛函变分, 得到

$$
\begin{aligned}
&\mathrm{d}\mathcal{L}(\boldsymbol{q}_0,\cdots,\boldsymbol{q}_n)(\delta\boldsymbol{q}_0,\cdots,\delta\boldsymbol{q}_n)\\
&=\sum_{k=0}^{n-1}\left[\frac{\partial}{\partial\boldsymbol{q}_1}L(\boldsymbol{q}_{k+1},\boldsymbol{q}_k)\delta\boldsymbol{q}_{k+1}+\frac{\partial}{\partial\boldsymbol{q}_0}L(\boldsymbol{q}_{k+1},\boldsymbol{q}_k)\delta\boldsymbol{q}_k\right]\\
&=\sum_{k=1}^{n}\frac{\partial}{\partial\boldsymbol{q}_1}L(\boldsymbol{q}_k,\boldsymbol{q}_{k-1})\delta\boldsymbol{q}_k+\sum_{k=0}^{n-1}\frac{\partial}{\partial\boldsymbol{q}_0}L(\boldsymbol{q}_{k+1},\boldsymbol{q}_k)\delta\boldsymbol{q}_k\\
&=\sum_{k=1}^{n-1}\left[\frac{\partial}{\partial\boldsymbol{q}_1}L(\boldsymbol{q}_k,\boldsymbol{q}_{k-1})+\frac{\partial}{\partial\boldsymbol{q}_0}L(\boldsymbol{q}_{k+1},\boldsymbol{q}_k)\right]\delta\boldsymbol{q}_k\\
&\quad+\frac{\partial}{\partial\boldsymbol{q}_0}L(\boldsymbol{q}_1,\boldsymbol{q}_0)\delta\boldsymbol{q}_0+\frac{\partial}{\partial\boldsymbol{q}_1}L(\boldsymbol{q}_n,\boldsymbol{q}_{n-1})\delta\boldsymbol{q}_n.
\end{aligned}\tag{4.3.2}
$$

固定端点, 得到离散的 Euler-Lagrange 方程组

$$
\frac{\partial}{\partial\boldsymbol{q}_1}L(\boldsymbol{q}_1,\boldsymbol{q}_0)+\frac{\partial}{\partial\boldsymbol{q}_0}L(\boldsymbol{q}_2,\boldsymbol{q}_1)=\mathbf{0},\tag{4.3.3}
$$

其对应离散的解流 $F_d: M\times M\to M\times M$:

$$
F_d(\boldsymbol{q}_1,\boldsymbol{q}_0)=(\boldsymbol{q}_2,\boldsymbol{q}_1).
$$

如何证明这个离散的流是辛的, 经典的处理就是: 定义离散的 Legendre 变换

$$
\begin{aligned}
&\mathbb{F}L: M\times M\to T^*M,\\
&(\boldsymbol{q}_{k+1},\boldsymbol{q}_k)\mapsto\left(\boldsymbol{q}_k,\frac{\partial}{\partial\boldsymbol{q}_0}L(\boldsymbol{q}_{k+1},\boldsymbol{q}_k)\right),
\end{aligned}
$$

如果又定义了离散的典则 1-形式 θ, 则通过上面定义的离散的 Legendre 变换拉回定义离散的 Hamilton 1-形式, 进而得到离散的 Hamilton 2-形式

$$
\omega=\mathbb{F}L^*\omega=\mathbb{F}L^*(-\mathrm{d}\theta)=-\mathbb{F}L^*(\mathrm{d}\theta).\tag{4.3.4}
$$

对于 Birkhoff 系统对应的 Pfaff 作用泛函问题, 这里暂不讨论它的 Legendre 变换. 下面我们看 Marsden 是如何应用 Veselov 离散方法的[WM97].

直接从离散的变分原理可以得到 $M\times M$ 上的 Lagrange 2-形式. 以变分 (4.3.2) 中的两个边界项定义两个 $M\times M$ 上的离散 Lagrange

1-形式:

$$\theta_L^-(\boldsymbol{q}_k,\boldsymbol{q}_{k-1})(\delta\boldsymbol{q}_k,\delta\boldsymbol{q}_{k-1}) \triangleq \frac{\partial}{\partial \boldsymbol{q}_0}L(\boldsymbol{q}_k,\boldsymbol{q}_{k-1})\delta\boldsymbol{q}_{k-1}, \tag{4.3.5}$$

$$\theta_L^+(\boldsymbol{q}_k,\boldsymbol{q}_{k-1})(\delta\boldsymbol{q}_k,\delta\boldsymbol{q}_{k-1}) \triangleq \frac{\partial}{\partial \boldsymbol{q}_1}L(\boldsymbol{q}_k,\boldsymbol{q}_{k-1})\delta\boldsymbol{q}_k. \tag{4.3.6}$$

将 θ_L^- 和 θ_L^+ 视做 Lagrange 1-形式的离散近似. 把离散变分限制在离散的变分解空间上, 则变分 (4.3.2) 导出

$$\mathrm{d}\mathcal{L} = \theta_L^- + F_d^*\theta_L^+, \tag{4.3.7}$$

进而得到辛流公式

$$F_d^*\omega_L = \omega_L, \quad \omega_L = \mathrm{d}\theta_L^- = -\mathrm{d}\theta_L^+. \tag{4.3.8}$$

这个结果和经典 Veselov 离散辛形式是一致的.

下面我们讨论 Pfaff 作用泛函的离散变分原理. 和连续的情形一样, 我们对 Pfaff 作用泛函在标准的相空间上作离散变分, 定义 Lagrange 1-形式 $\theta_L = \boldsymbol{F}\mathrm{d}\boldsymbol{z}$ 的两个离散形式为

$$\begin{aligned}
\theta_L^-(\boldsymbol{z}_k,\boldsymbol{z}_{k-1}) &\triangleq \left(\frac{\partial}{\partial z_0}\boldsymbol{F}(\boldsymbol{z}_k,\boldsymbol{z}_{k-1}) - \frac{\partial}{\partial z_0}B(\boldsymbol{z}_k,\boldsymbol{z}_{k-1})\right)\delta\boldsymbol{z}_{k-1},\\
\theta_L^+(\boldsymbol{z}_k,\boldsymbol{z}_{k-1}) &\triangleq \left(\frac{\partial}{\partial z_1}\boldsymbol{F}(\boldsymbol{z}_k,\boldsymbol{z}_{k-1}) - \frac{\partial}{\partial z_1}B(\boldsymbol{z}_k,\boldsymbol{z}_{k-1})\right)\delta\boldsymbol{z}_{k-1},
\end{aligned} \tag{4.3.9}$$

最后得到和 Marsden 一样的离散变分. 与 Marsden 的方法有所不同, 我们在增广空间 $\widehat{Z}$ 上对变分进行如下离散: $\widehat{Z}\times\widehat{Z}$ 作为切丛 $T\widehat{Z}$ 的离散空间, 设离散的时间步长为 τ (也可以是变步长, 并不影响后面的离散变分). 定义离散的 Pfaff 作用泛函为

$$\mathcal{L}(\widehat{\boldsymbol{z}}) = \sum_{k=1}^{n}\mathfrak{F}(\boldsymbol{z}_k,\boldsymbol{z}_{k-1},s_k,s_{k-1}) - \mathfrak{B}(\boldsymbol{z}_k,\boldsymbol{z}_{k-1},s_k,s_{k-1}), \tag{4.3.10}$$

其中 $\mathfrak{F}$ 和 $\mathfrak{B}$ 是 $\widehat{Z}\times\widehat{Z}$ 上的光滑映射.

计算 $\mathcal{L}$ 的离散变分:

$$
\begin{aligned}
\mathrm{d}\mathcal{L} = & \sum_{k=1}^{n}\left[\frac{\partial}{\partial z_1}\mathfrak{F}(z_k, z_{k-1}, s_k, s_{k-1})\delta z_k\right. \\
& +\frac{\partial}{\partial z_0}\mathfrak{F}(z_k, z_{k-1}, s_k, s_{k-1})\delta z_{k-1} \\
& \left.+\frac{\partial}{\partial s_1}\mathfrak{F}(z_k, z_{k-1}, s_k, s_{k-1})\delta s_k + \frac{\partial}{\partial s_0}\mathfrak{F}(z_k, z_{k-1}, s_k, s_{k-1})\delta s_{k-1}\right] \\
& -\sum_{k=1}^{n}\left[\frac{\partial}{\partial z_1}\mathfrak{B}(z_k, z_{k-1}, s_k, s_{k-1})\delta z_k\right. \\
& +\frac{\partial}{\partial z_0}\mathfrak{B}(z_k, z_{k-1}, s_k, s_{k-1})\delta z_{k-1} - \frac{\partial}{\partial s_1}\mathfrak{B}(z_k, z_{k-1}, s_k, s_{k-1})\delta s_k \\
& \left.-\frac{\partial}{\partial s_0}\mathfrak{B}(z_k, z_{k-1}, s_k, s_{k-1})\delta s_{k-1}\right] \\
= & \sum_{k=1}^{n}\left[\frac{\partial}{\partial z_1}\mathfrak{F}(z_k, z_{k-1}, s_k, s_{k-1})\delta z_k + \frac{\partial}{\partial s_1}\mathfrak{F}(z_k, z_{k-1}, s_k, s_{k-1})\delta s_k\right. \\
& \left.-\frac{\partial}{\partial z_1}\mathfrak{B}(z_k, z_{k-1}, s_k, s_{k-1})\delta z_k - \frac{\partial}{\partial s_1}\mathfrak{B}(z_k, z_{k-1}, s_k, s_{k-1})\delta s_k\right] \\
& +\sum_{k=0}^{n-1}\left[\frac{\partial}{\partial z_0}\mathfrak{F}(z_{k+1}, z_k, s_{k+1}, s_k)\delta z_k + \frac{\partial}{\partial s_0}\mathfrak{F}(z_{k+1}, z_k, s_{k+1}, s_k)\delta s_k\right. \\
& \left.-\frac{\partial}{\partial z_0}\mathfrak{B}(z_{k+1}, z_k, s_{k+1}, s_k)\delta z_k - \frac{\partial}{\partial s_0}\mathfrak{B}(z_{k+1}, z_k, s_{k+1}, s_k)\delta s_k\right] \\
= & \sum_{k=1}^{n-1}\left[\frac{\partial}{\partial z_1}\mathfrak{F}(z_k, z_{k-1}, s_k, s_{k-1}) + \frac{\partial}{\partial z_0}\mathfrak{F}(z_{k+1}, z_k, s_k, s_{k-1})\right. \\
& \left.-\frac{\partial}{\partial z_1}\mathfrak{B}(z_k, z_{k-1}, s_k, s_{k-1}) - \frac{\partial}{\partial z_0}\mathfrak{B}(z_{k+1}, z_k, s_{k+1}, s_k)\right]\delta z_k \\
& +\sum_{k=1}^{n-1}\left[\frac{\partial}{\partial s_1}\mathfrak{F}(z_k, z_{k-1}, s_k, s_{k-1}) + \frac{\partial}{\partial s_0}\mathfrak{F}(z_{k+1}, z_k, s_{k+1}, s_k)\right. \\
& \left.-\frac{\partial}{\partial s_1}\mathfrak{B}(z_k, z_{k-1}, s_k, s_{k-1}) - \frac{\partial}{\partial s_0}\mathfrak{B}(z_{k+1}, z_k, s_{k+1}, s_k)\right]\delta s_k \\
& +\left[\frac{\partial}{\partial z_1}\mathfrak{F}(z_n, z_{n-1}, s_n, s_{n-1}) - \frac{\partial}{\partial z_1}\mathfrak{B}(z_n, z_{n-1}, s_n, s_{n-1})\right]\delta z_n
\end{aligned}
$$

$$+\left[\frac{\partial}{\partial s_1}\mathfrak{F}(z_n,z_{n-1},s_n,s_{n-1})-\frac{\partial}{\partial s_1}\mathfrak{B}(z_n,z_{n-1},s_n,s_{n-1})\right]\delta s_n$$
$$+\left[\frac{\partial}{\partial z_0}\mathfrak{F}(z_1,z_0,s_1,s_0)-\frac{\partial}{\partial z_0}\mathfrak{B}(z_1,z_0,s_1,s_0)\right]\delta z_0$$
$$+\left[\frac{\partial}{\partial s_0}\mathfrak{F}(z_1,z_0,s_1,s_0)-\frac{\partial}{\partial s_0}\mathfrak{B}(z_1,z_0,s_1,s_0)\right]\delta s_0. \tag{4.3.11}$$

固定端点, 得到离散 Euler-Lagrange 方程

$$\frac{\partial}{\partial z_1}\mathfrak{F}(z_k,z_{k-1},s_k,s_{k-1})+\frac{\partial}{\partial z_0}\mathfrak{F}(z_{k+1},z_k,s_k,s_{k-1})$$
$$-\frac{\partial}{\partial z_1}\mathfrak{B}(z_k,z_{k-1},s_k,s_{k-1})-\frac{\partial}{\partial z_0}\mathfrak{B}(z_{k+1},z_k,s_{k+1},s_k)=\mathbf{0}. \tag{4.3.12}$$

定义离散的能量

$$E_d=\frac{\partial}{\partial s_1}\mathfrak{F}(z_k,z_{k-1},s_k,s_{k-1})-\frac{\partial}{\partial s_1}\mathfrak{B}(z_k,z_{k-1},s_k,s_{k-1})$$
$$=\frac{\partial}{\partial s_0}\mathfrak{F}(z_{k+1},z_k,s_{k+1},s_k)-\frac{\partial}{\partial s_0}\mathfrak{B}(z_{k+1},z_k,s_{k+1},s_k). \tag{4.3.13}$$

根据边界项定义一对离散 1-形式:

$$\theta_L^+=\left[\frac{\partial}{\partial z_1}\mathfrak{F}(z_k,z_{k-1},s_k,s_{k-1})-\frac{\partial}{\partial z_1}\mathfrak{B}(z_k,z_{k-1},s_k,s_{k-1})\right]\delta z_k$$
$$+\left[\frac{\partial}{\partial s_1}\mathfrak{F}(z_k,z_{k-1},s_k,s_{k-1})-\frac{\partial}{\partial s_1}\mathfrak{B}(z_k,z_{k-1},s_k,s_{k-1})\right]\delta s_k, \tag{4.3.14}$$

$$\theta_L^-=\left[\frac{\partial}{\partial z_0}\mathfrak{F}(z_k,z_{k-1},s_k,s_{k-1})-\frac{\partial}{\partial z_0}\mathfrak{B}(z_k,z_{k-1},s_k,s_{k-1})\right]\delta z_{k-1}$$
$$+\left[\frac{\partial}{\partial s_0}\mathfrak{F}(z_k,z_{k-1},s_k,s_{k-1})-\frac{\partial}{\partial s_0}\mathfrak{B}(z_k,z_{k-1},s_k,s_{k-1})\right]\delta s_{k-1}. \tag{4.3.15}$$

直接相加 θ_L^+ 和 θ_L^-, 得到

$$\theta_L^+ + \theta_L^- = \widehat{\mathrm{d}}L(\widehat{z}), \tag{4.3.16}$$

所以有 $\widehat{\mathrm{d}}\theta_L^+ + \widehat{\mathrm{d}}\theta_L^- = 0$. 定义

$$\widehat{\omega} = \widehat{\mathrm{d}}\theta_L^- = -\widehat{\mathrm{d}}\theta_L^+. \tag{4.3.17}$$

假设增广空间 $\widehat{Z}$ 上离散的 Lagrange 解流为

$$\widehat{F}_d : \widehat{Z} \times \widehat{Z} \to \widehat{Z} \times \widehat{Z},$$

$$(\boldsymbol{z}_{k-1}, \boldsymbol{z}_k, s_{k-1}, s_k) \mapsto (\boldsymbol{z}_k, \boldsymbol{z}_{k+1}, s_k, s_{k+1}),$$

那么它包含了原相空间上的相流

$$F_d : Z \times Z \to Z \times Z,$$

$$(\boldsymbol{z}_{k-1}, \boldsymbol{z}_k) \mapsto (\boldsymbol{z}_k, \boldsymbol{z}_{k+1}).$$

将变分限制在 $[t_k, t_{k+1}]$ 上, 得到

$$\widehat{\mathrm{d}}\mathfrak{L} = \widehat{\theta}^- + \widehat{F}_d^*\widehat{\theta}^+. \tag{4.3.18}$$

对上式外微分, 得到辛流公式

$$\widehat{F}_d^*\widehat{\omega} = \widehat{\omega}. \tag{4.3.19}$$

我们再次强调: 上面的变分过程中的外微分记号 d 表示在 Z 上变分, 而 $\widehat{\mathrm{d}}$ 表示在增广空间 $\widehat{Z}$ 上变分, 不同于直接在 Z 的相空间上变分. 如果令

$$\begin{aligned}
\theta^+ &= \left(\frac{\partial}{\partial \boldsymbol{z}_1}\mathfrak{F}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, s_k, s_{k-1}) - \frac{\partial}{\partial \boldsymbol{z}_1}\mathfrak{B}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, s_k, s_{k-1})\right)\delta \boldsymbol{z}_k,\\
\theta^- &= \left(\frac{\partial}{\partial \boldsymbol{z}_0}\mathfrak{F}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, s_k, s_{k-1}) - \frac{\partial}{\partial \boldsymbol{z}_0}\mathfrak{B}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, s_k, s_{k-1})\right)\delta \boldsymbol{z}_{k-1},
\end{aligned} \tag{4.3.20}$$

那么 $\theta^- + \theta^+ = \mathrm{d}L(\widehat{\boldsymbol{z}}) \Longrightarrow \mathrm{d}\theta^- + \mathrm{d}\theta^+ = 0$. 所以定义

$$\omega = \mathrm{d}\theta^- = -\mathrm{d}\theta^+. \tag{4.3.21}$$

令 $\mathrm{d}^s = \widehat{\mathrm{d}} - \mathrm{d}$. 因为

$$\theta_L^+ = \theta^+ + A^+, \quad \theta_L^- = \theta^- + A^-, \tag{4.3.22}$$

$$A^+ = \left(\frac{\partial}{\partial s_1}\mathfrak{F}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, s_k, s_{k-1}) - \frac{\partial}{\partial s_1}\mathfrak{B}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, s_k, s_{k-1})\right)\delta s_k,$$

$$A^- = \left(\frac{\partial}{\partial s_0}\mathfrak{F}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, s_k, s_{k-1}) - \frac{\partial}{\partial s_0}\mathfrak{B}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, s_k, s_{k-1})\right)\delta s_{k-1},$$

以及

$$\widehat{\mathrm{d}}\theta_L^+ + \widehat{\mathrm{d}}\theta_L^- = 0 \quad 和 \quad \mathrm{d}\theta_L^+ + \mathrm{d}\theta_L^- = 0,$$

所以有

$$\widehat{\mathrm{d}}A^+ + \mathrm{d}^s\theta^+ = -\widehat{\mathrm{d}}A^- - \mathrm{d}^s\theta^-. \tag{4.3.23}$$

由辛流公式 (4.3.19), 将 $s_k = t_k$ 代入 ω, 得到

$$\omega + \widehat{\mathrm{d}}A^- + \mathrm{d}^s\theta^- = \widehat{\omega} = F_d^*(\omega) - \widehat{F}_d^*(\widehat{\mathrm{d}}A^+ + \mathrm{d}^s\theta^+), \tag{4.3.24}$$

所以我们得到标准相空间上离散的约束辛流公式

$$F_d^*\omega = \omega, \quad -\widehat{F}_d^*(\widehat{\mathrm{d}}A^+ + \mathrm{d}^s\theta^+) = \widehat{\mathrm{d}}A^- + \mathrm{d}^s\theta^-. \tag{4.3.25}$$

上面的离散约束条件表示 (4.3.13) 式定义的离散能量满足的相流公式, 我们认为它是连续约束条件的离散形式. 如果直接在标准相空间上考虑离散变分并定义离散 1-形式和 2-形式, 我们同样可以得到辛流公式 $F_d^*\omega = \omega$, 但是无法得到约束条件的离散形式.

再对照全变分的离散情形. 用 $Z \times Z \times \mathbf{R}$ 作为切丛 TZ 的离散空间, 设离散的变时间步长为 h_k, $Z \times Z \times \mathbf{R}$ 中的点 $(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, h_k)$ 对应切向量 $(\boldsymbol{z}_k - \boldsymbol{z}_{k-1})/h_k$, 离散的 Lagrange 函数是一个 $Z \times Z \times \mathbf{R}$ 到 $\mathbf{R}$ 的光滑映射. 我们记离散 Pfaff 作用泛函为

$$\mathfrak{L}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, h_k) = \sum_{k=1}^{n} h_k \left[\mathfrak{F}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, h_k) - \mathfrak{B}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, h_k)\right]. \tag{4.3.26}$$

对 $\mathfrak{L}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, h_k)$ 全变分, 得到

$$\begin{aligned}
&\mathrm{d}\mathfrak{L}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, h_k) \\
&\quad = \sum_{k=1}^{n}\left[h_k\frac{\partial}{\partial \boldsymbol{z}_1}\mathfrak{F}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, h_k) - \frac{\partial}{\partial \boldsymbol{z}_1}\mathfrak{B}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, h_k)\right]\delta \boldsymbol{z}_k \\
&\quad\quad + \sum_{k=1}^{n}\left[h_k\frac{\partial}{\partial \boldsymbol{z}_0}\mathfrak{F}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, h_k) - \frac{\partial}{\partial \boldsymbol{z}_0}\mathfrak{B}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, h_k)\right]\delta \boldsymbol{z}_{k-1}
\end{aligned}$$

$$
\begin{aligned}
&+\sum_{k=1}^{n}\frac{\partial}{\partial h_k}[h_k\mathfrak{F}(z_k,z_{k-1},h_k)-h_k\mathfrak{B}(z_k,z_{k-1},h_k)]\delta h_k\\
=&\sum_{k=1}^{n-1}\left[h_k\frac{\partial}{\partial z_1}\mathfrak{F}(z_k,z_{k-1},h_k)+h_{k+1}\frac{\partial}{\partial z_0}\mathfrak{F}(z_{k+1},z_k,h_{k+1})\right.\\
&\left.-h_k\frac{\partial}{\partial z_1}\mathfrak{B}(z_k,z_{k-1},h_k)-h_{k+1}\frac{\partial}{\partial z_0}\mathfrak{B}(z_{k+1},z_k,h_{k+1})\right]\delta z_k\\
&+\sum_{k=2}^{n-1}\frac{\partial}{\partial h_k}\left[h_k\mathfrak{F}(z_k,z_{k-1},h_k)-h_k\mathfrak{B}(z_k,z_{k-1},h_k)\right]\delta h_k\\
&+\left[h_n\frac{\partial}{\partial z_1}\mathfrak{F}(z_n,z_{n-1},h_n)-h_n\frac{\partial}{\partial z_1}\mathfrak{B}(z_n,z_{n-1},h_n)\right]\delta z_n\\
&+\frac{\partial}{\partial h_n}\left[h_n\mathfrak{F}(z_n,z_{n-1},h_n)-h_n\mathfrak{B}(z_n,z_{n-1},h_n)\right]\delta h_n\\
&+\left[h_1\frac{\partial}{\partial z_0}\mathfrak{F}(z_1,z_0,h_1))-h_1\frac{\partial}{\partial z_0}\mathfrak{B}(z_1,z_0,h_1)\right]\delta z_0\\
&+\frac{\partial}{\partial h_1}\left[h_1\mathfrak{F}(z_1,z_0,h_1)-h_1\mathfrak{B}(z_1,z_0,h_1)\right]\delta h_1. \qquad (4.3.27)
\end{aligned}
$$

同样, 固定端点, 得到离散 Euler-Lagrange 方程

$$
\begin{aligned}
&h_k\frac{\partial}{\partial z_1}\mathfrak{F}(z_k,z_{k-1},h_k)+h_{k+1}\frac{\partial}{\partial z_0}\mathfrak{F}(z_{k+1},z_k,h_{k+1})\\
&-h_k\frac{\partial}{\partial z_1}\mathfrak{B}(z_k,z_{k-1},h_k)-h_{k+1}\frac{\partial}{\partial z_0}\mathfrak{B}(z_{k+1},z_k,h_{k+1})=0. \qquad (4.3.28)
\end{aligned}
$$

它和离散 Euler-Lagrange 方程 (4.3.12) 是完全等价的. 记

$$
\mathbb{E}(z_k,z_{k-1},h_k)=\frac{\partial}{\partial h_k}\left[h_k\mathfrak{F}(z_k,z_{k-1},h_k)-h_k\mathfrak{B}(z_k,z_{k-1},h_k)\right]. \qquad (4.3.29)
$$

令

$$
\mathbb{E}(z_k,z_{k-1},h_k)=\mathbb{E}(z_{k+1},z_k,h_{k+1}). \qquad (4.3.30)
$$

由边界项定义 θ_L^+ 和 θ_L^-:

$$
\begin{aligned}
\theta_L^+=&\left[h_k\frac{\partial}{\partial z_1}\mathfrak{F}(z_k,z_{k-1},h_k)-h_k\frac{\partial}{\partial z_1}\mathfrak{B}(z_k,z_{k-1},h_k)\right]\delta z_k\\
&+\mathbb{E}(z_k,z_{k-1},h_k)\delta h_k,\\
\theta_L^-=&\left[h_k\frac{\partial}{\partial z_0}\mathfrak{F}(z_k,z_{k-1},h_k)-h_k\frac{\partial}{\partial z_0}\mathfrak{B}(z_k,z_{k-1},h_k)\right]\delta z_{k-1}\\
&+\mathbb{E}(z_k,z_{k-1},h_k)\delta h_k,
\end{aligned} \qquad (4.3.31)
$$

令 $\theta^+ = \theta_L^+ - \mathbb{E}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, h_k)\delta h_k$, $\theta^- = \theta_L^- - \mathbb{E}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, h_k)\delta h_k$, 则

$$\theta^- + \theta^+ = \mathrm{d}(h_k(\mathfrak{F}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, h_k) - \mathfrak{B}(\boldsymbol{z}_k, \boldsymbol{z}_{k-1}, h_k))). \tag{4.3.32}$$

离散的 Lagrange 流仍记为 F_d, 令 $\Omega = -\mathrm{d}\theta_L^+$, $\omega = -\mathrm{d}\theta^+$, 则可以推出相流若是包括时间 t 的流, 那么有公式

$$F_d^* \Omega = \omega. \tag{4.3.33}$$

再进一步, 在局部坐标上我们也可以推出上式等价于离散的标准辛流公式加上离散约束条件.

§4.4　经典场论的变分原理

我们首先介绍几个基本概念, 包括构形丛, Jet 丛和对偶 Jet 丛等, 它们分别是 Lagrange 力学中构形空间、切空间和余切空间在场论中的推广.

定义 4.4.1　假设 X 是一个 $n+1$ 维时–空微分流形, 其局部坐标为 $\boldsymbol{x} = (x^0, x^1, \cdots, x^n)$, 其中 x^0 表示时间方向, $x^\nu(\nu = 1, \cdots, n)$ 表示空间方向. Y 是一个目标空间, 它的局部坐标为 $\boldsymbol{y} = (y^1, \cdots, y^N)$. 令 $\pi_{XY}: Y \to X$ 表示 X 上的纤维, 一个协变的**构形丛**就是纤维丛 (Y, X, π_{XY}), 简记为 π_{XY}. 设 $\pi_{XY}^{-1}: \boldsymbol{x} \mapsto Y_{\boldsymbol{x}}$, 即 $Y_{\boldsymbol{x}}$ 是 π_{XY} 在点 $\boldsymbol{x}$ 处的纤维, 又设线性映射 $\gamma: T_{\boldsymbol{x}}X \to T_{\boldsymbol{y}}Y$ $(\boldsymbol{y} \in Y_{\boldsymbol{x}})$ 满足

$$T_{\pi_{XY}} \circ \gamma = \mathrm{id}, \tag{4.4.1}$$

即 $T_{\pi_{XY}} \circ \gamma$ 是 $T_{\boldsymbol{x}}X$ 上的恒等映射. 这样的映射组成仿射丛 AFF $(T_{\boldsymbol{x}}X, T_{\boldsymbol{y}}Y)$, 即 $\boldsymbol{y}$ 处的纤维. 由这样的纤维构成的丛定义为 Y 上的**一阶 Jet 丛**, 记为 $J^1(Y)$.

假设映射 $\phi: X \to Y$ 是构形丛 π_{XY} 的一个截面, 它的切映射用 $T_{\boldsymbol{x}}\phi$ 表示, 则 $T_{\boldsymbol{x}}\phi \in J^1(Y)_{\phi(\boldsymbol{x})}$, 而映射: $\boldsymbol{x} \mapsto T_{\boldsymbol{x}}\phi$ 是 $J^1(Y)$ 的一个截面, 也可以看做 X 上的一个丛, 记做 $j^1(\phi)$, 称为 ϕ 的**一阶 Jet**. $j^1(\phi)$ 用局部坐标表示为

$$x^\nu \mapsto (x^\mu, \phi^A(\boldsymbol{x}), \partial_\nu \phi^A(\boldsymbol{x})), \quad \partial_\nu = \frac{\partial}{\partial x^\nu}. \tag{4.4.2}$$

定义 4.4.2 设 $\Lambda^{n+1}(X)$ 是 X 上的 $(n+1)$-形式集合, 由映射 $J^1(Y)_{\boldsymbol{y}} \to \Lambda^{n+1}(X)_{\boldsymbol{x}}$ $(\boldsymbol{y} \in Y_{\boldsymbol{x}})$ 组成的向量丛称为 Y 上的一阶**对偶 Jet 丛**, 记为 $J^1(Y)^*$. 覆盖 π_{XY} 的映射 $J^1(Y) \to \Lambda^{n+1}(X)$, 称为 $J^1(Y)^*$ 的一个**截面**.

场论变分的 Lagrange 函数 L 在坐标意义下表示为

$$L(\gamma) = L(x^\mu, y^A, y^A_\nu)\mathrm{d}^{n+1}\boldsymbol{x}, \tag{4.4.3}$$

由 $L(\gamma)$ 的一阶 Taylor 近似定义一个从 Jet 丛 $J^1(Y)$ 到其对偶 Jet 丛的保纤维映射, 即协变的 Legendre 变换, 通过 Legendre 变换将对偶 Jet 丛 $J^1(Y)^*$ 上的典则 $(n+1)$-形式 θ 拉回得到 $J^1(Y)$ 上的 Cartan 形式. 另外, 对由 $L(\gamma)$ 定义的 $\pounds(\gamma)$ 进行场变分直接得到其多辛结构公式. 下面对 Marsden 多辛场论的变分方法的结果做简单介绍[MPS98].

定义 4.4.3 令 U 是具有光滑闭边界的光滑微分流形. 定义

$$\mathcal{C}^\infty = \left\{\phi : U \to Y | \pi_{XY} \circ \phi : U \to X \text{ 是嵌入映射}\right\}, \tag{4.4.4}$$

称之为**光滑映射集**, 简记为 $\mathcal{C}$.

对每个 $\phi \in \mathcal{C}^\infty$, 记 $\phi_X = \pi_{XY} \circ \phi$, 则 $\phi_X : U \to U_X \subset X$ 是微分同胚.

不妨假设 $\mathcal{C}$ 是微分流形, 它的切空间为 $T_\phi\mathcal{C}$. 再假设 G 是由 π_{XY} 丛自同构 η_Y 构成的 Lie 群, η_Y 覆盖微分同胚 η_X, G 的 Lie 代数为 $\mathcal{G}$. 定义作用 $\Phi : G \times \mathcal{C} \to \mathcal{C}$ 为

$$\Phi(\eta_Y, \phi) = \eta_Y \circ \phi \circ \eta_X^{-1}. \tag{4.4.5}$$

我们对所有的截面 ϕ 定义作用函数:

$$S(\phi) = \int_{U_X} L(j^1(\phi \circ \phi_X^{-1})). \tag{4.4.6}$$

令 $\lambda \mapsto \eta_Y^\lambda$ 是 G 中的任意光滑路径, $\eta_Y^0 = e$ (e 为 G 的单位元), $V \in T_\phi\mathcal{C}$ 由下式给出:

$$V = \frac{\mathrm{d}}{\mathrm{d}\lambda}\Phi(\eta_Y^\lambda, \phi)\Big|_{\lambda=0}. \tag{4.4.7}$$

对作用函数 (4.4.6) 作场变分, 简单地讲, 就是将 ϕ 做个扰动, 即 $\Phi(\eta_Y^\lambda,\phi)$, 然后寻求临界截面使得

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}S(\Phi(\eta_Y^\lambda,\phi))\Big|_{\lambda=0}=0. \tag{4.4.8}$$

这和 Lagrange 力学变分有些类似. 以 (4.4.7) 式定义向量场的一阶无穷小生成子, 令

$$V=V^h+V^v,\quad V^h=T(\phi\circ\phi_X^{-1})V_X,\quad V^v=V-V^h, \tag{4.4.9}$$

其中 $V_X=\dfrac{\mathrm{d}}{\mathrm{d}\lambda}\eta_Y^\lambda\circ\phi\Big|_{\lambda=0}$, 则

$$\begin{aligned}\mathrm{d}S_\phi\cdot V&=\frac{\mathrm{d}}{\mathrm{d}\lambda}\int_{\eta_X^\lambda\circ\phi_X(U)}L(j^1(\Phi(\eta_Y^\lambda,\phi)))\Big|_{\lambda=0}\\&=\frac{\mathrm{d}}{\mathrm{d}\lambda}\int_{\phi_X(U)}j^1(\phi\circ\phi_X^{-1})^*j^1(\eta_Y^\lambda)^*\theta_L\Big|_{\lambda=0}\\&=\int_{U_X}\frac{\mathrm{d}}{\mathrm{d}\lambda}L(j^1(\Phi(\eta_Y^\lambda,\phi)))\Big|_{\lambda=0}+\int_{U_X}\mathrm{L}_{V_X}(L(j^1(\phi\circ\phi_X^{-1}))),\end{aligned} \tag{4.4.10}$$

这里应用了 $L(z)=z^*\theta_L$ 以及

$$j^1(\eta_Y\circ\phi\circ\phi_X^{-1}\circ\eta_X^{-1})=j^1(\eta_Y)\circ j^1(\phi\circ\phi_X^{-1})\circ\eta_X^{-1}. \tag{4.4.11}$$

θ_L 就是 Cartan $(n+1)$-形式. 特别注意 (4.4.10) 式中的 L_{V_X} 是指关于 V_X 的李导数.

由上面的推导可得到下面的引理:

引理 4.4.1 对于任意的 $V\in T_\phi\mathcal{C}$, 有

$$\mathrm{d}S_\phi\cdot V^h=\int_{\partial U_X}i_{V_X}L(j^1(\phi\circ\phi_X^{-1})) \tag{4.4.12}$$

和

$$\mathrm{d}S_\phi\cdot V^v=\int_{U_X}L(j^1(\Phi(\eta_Y^\lambda,\phi))). \tag{4.4.13}$$

定理 4.4.1 给定一个光滑的 Lagrange 函数 $L:J^1(Y)\to\Lambda^{n+1}(X)$, 存在唯一的 C^∞ 光滑截面 $D_{EL}L:Y''\to\Lambda^{n+1}(X)\otimes T^*Y$ 以及唯一的

微分形式 $\theta_L \in \Lambda^{n+1}(J^1(Y))$, 使得对于任意的 $V \in T_\phi \mathcal{C}$ 和 X 的任意开子集 U_X ($\overline{U}_X \cap \partial X = \varnothing$), 成立

$$\mathrm{d}S_\phi \cdot V = \int_{U_X} D_{EL}L(j^2(\phi \circ \phi_X^{-1})) \cdot V + \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^*(i_{j^1(V)}\theta_L), \tag{4.4.14}$$

且在 U_X 上有

$$D_{EL}L(j^2(\phi \circ \phi_X^{-1})) \cdot V = j^1(\phi \circ \phi_X^{-1})^*(i_{j^1(V)}\omega_L), \tag{4.4.15}$$

即 Y'' 上的 Euler-Lagrange 算子作用在 L 上得到的 $D_{EL}L$ 可用局部坐标表示为

$$\begin{aligned} &D_{EL}L(j^2(\phi \circ \phi_X^{-1})) \\ &= \left[\frac{\partial L}{\partial y^A}(j^1(\phi \circ \phi_X^{-1})) - \frac{\partial^2 L}{\partial x^\mu \partial v_\mu^A}(j^1(\phi \circ \phi_X^{-1}))\right. \\ &\quad - \frac{\partial^2 L}{\partial y^B \partial v_\mu^A}(j^1(\phi \circ \phi_X^{-1})) \cdot (\phi \circ \phi_X^{-1})^B_{,\mu} \\ &\quad \left. - \frac{\partial^2 L}{\partial v_\nu^B \partial v_\mu^A}(j^1(\phi \circ \phi_X^{-1})) \cdot (\phi \circ \phi_X^{-1})^B_{,\mu\nu}\right] \mathrm{d}y^A \wedge \mathrm{d}^{n+1}\boldsymbol{x}, \end{aligned} \tag{4.4.16}$$

其中 $\omega_L = \mathrm{d}\theta_L$, θ_L 即 Cartan 形式, 其坐标表示为

$$\theta_L = \frac{\partial L}{\partial v_\mu^A}\mathrm{d}y^A \wedge \mathrm{d}^n x_\mu + \left(L - \frac{\partial L}{\partial v_\mu^A}v_\mu^A\right)\mathrm{d}^{n+1}\boldsymbol{x}, \tag{4.4.17}$$

$v_\mu^A \doteq \dfrac{\partial y^A}{\partial x_\mu}$, $(\phi \circ \phi_X^{-1})^B_{,\mu}$ 表示第 B 个分量关于 x_μ 求偏导数, 而 $(\phi \circ \phi_X^{-1})^B_{,\mu\nu}$ 表示第 B 个分量关于 x_μ 和 x_ν 求二阶导数, 这里 $A, B = 1, \cdots, N$; $\mu, \nu = 0, 1, \cdots, n$.

定理证明参见文献 [MPS98].

§4.5 高维 Pfaff 作用泛函的场论变分

本节我们讨论经典场论变分在 Pfaff 作用泛函问题上的推广. $n+1$ 维空间上的 Pfaff 作用泛函定义为

$$\mathcal{L}(\boldsymbol{y}, \boldsymbol{x}, t) = \int_X \left[F_\nu^\mu(\boldsymbol{y}, \boldsymbol{x}, t) z_{x_\mu}^\nu - B(\boldsymbol{y}, \boldsymbol{x})\right] \mathrm{d}^n\boldsymbol{x} \wedge \mathrm{d}t. \tag{4.5.1}$$

参照定理 4.4.1, 我们可以很容易证明以下结论:

定理 4.5.1 假设 Pfaff 作用泛函 (4.5.1) 给定的一阶 Lagrange 函数 $L: J^1(Y) \to \Lambda^{n+1}(X)$ 是光滑的, 那么存在唯一的 C^∞ 光滑截面 $D_{EL}L: Y' \to \Lambda^{n+1}(X) \otimes T^*Y$ 以及唯一的微分形式 $\theta_L \in \Lambda^{n+1}(J^1(Y))$, 使得对于任意的 $V \in T_\phi \mathcal{C}$ 和 X 的任意开子集 U_X ($\overline{U}_X \cap \partial X = \varnothing$), 成立

$$\mathrm{d}S_\phi \cdot V = \int_{U_X} D_{EL}L(j^1(\phi \circ \phi_X^{-1})) \cdot V + \int_{\partial U_X} j^1(\phi \circ \phi_X^{-1})^*(i_{j^1(V)}\theta_L), \tag{4.5.2}$$

且在 U_X 上有

$$D_{EL}L(j^1(\phi \circ \phi_X^{-1})) \cdot V = j^1(\phi \circ \phi_X^{-1})^*(i_{j^1(V)}\omega_L), \tag{4.5.3}$$

即 Y' 上的 Euler-Lagrange 算子作用在 L 上得到的 $D_{EL}L$ 可用局部坐标表示为

$$\begin{aligned} &D_{EL}L(j^1(\phi \circ \phi_X^{-1})) \\ &\quad = \left[\frac{\partial L}{\partial y^A}(j^1(\phi \circ \phi_X^{-1})) - \frac{\partial^2 L}{\partial x^\mu \partial v_\mu^A}(j^1(\phi \circ \phi_X^{-1}))\right. \\ &\qquad \left. - \frac{\partial^2 L}{\partial y^B \partial v_\mu^A}(j^1(\phi \circ \phi_X^{-1})) \cdot (\phi \circ \phi_X^{-1})^B_{,\mu}\right] \mathrm{d}y^A \wedge \mathrm{d}^n \boldsymbol{x} \wedge \mathrm{d}t, \end{aligned} \tag{4.5.4}$$

其中 $\omega_L = \mathrm{d}\theta_L$, θ_L 即 Cartan 形式, 其坐标表示为

$$\theta_L = \frac{\partial L}{\partial v_\mu^A}\mathrm{d}y^A \wedge \mathrm{d}^n x_\mu + \left(L - \frac{\partial L}{\partial v_\mu^A} v_\mu^A\right)\mathrm{d}^n \boldsymbol{x} \wedge \mathrm{d}t. \tag{4.5.5}$$

和 Lagrange 力学变分一样, 我们发现这样得到的几何结构只是刻画了 Pfaff 作用泛函问题的部分性质, 所以我们需要对以上定理做一些修改. 为了说明起来简单, 我们以 $1+1$ 维空间为例. 引进新的变量 $\boldsymbol{s} = (s_1, s_2)$, $s_1 = t$, $s_2 = x$. 令 $X = \{(x,t) | x, t \in \mathbf{R}\}$, $Y = \widehat{Z} = \{(\boldsymbol{z}, \boldsymbol{s}) | \boldsymbol{z} \in \mathbf{R}^N\}$. 若 $U \subset X$ 是具有光滑闭边界的光滑流形, 则对应的 $\mathcal{L}$ 的作用函数可以写为

$$S(\phi, x, t) = \int_U [\mathfrak{F}(\boldsymbol{s}, \phi, \phi_t) + \mathfrak{G}(\boldsymbol{s}, \phi, \phi_x) - \mathfrak{B}(\boldsymbol{s}, \phi, s_t, s_x)]\, \mathrm{d}x \wedge \mathrm{d}t. \tag{4.5.6}$$

其中 $\mathfrak{F}, \mathfrak{G}$ 和 $\mathfrak{B}$ 都是 $J^1(Y)$ 上的光滑映射. 设 $\widehat{G}$ 是 π_{XY} 从自同构 η_Y 的 Lie 群, 其中 η_Y 覆盖微分同胚 η_X. 令 $\lambda \mapsto \eta_Y^\lambda$ 是 $\widehat{G}$ 中任意使得

$\eta_Y^0 = \widehat{e}$ 的光滑路径, 那么伴随有一个覆盖微分同胚 η_X 的 π_{XZ} 从自同构的 Lie 子群 G, 以及其中的任意光滑路径 $\lambda \mapsto \eta_Z^\lambda$, 并且 $\eta_Z^0 = e$. 当 λ 充分小时, η_X^λ 是局部可逆的, 则 $\widehat{G}$ 对 $\widehat{\phi} \in \widehat{\mathcal{C}}$ 的群作用可以描述为

$$\widetilde{\widehat{\phi}} = \eta_Y^\lambda \circ \widehat{\phi} \circ (\eta_X^\lambda)^{-1}, \tag{4.5.7}$$

它包含了 $\phi \in \mathcal{C}$ 上的 Lie 子群作用 $G \times \mathcal{C} \to \mathcal{C}$:

$$\widetilde{\phi} = \eta_Z^\lambda \circ \phi \circ (\eta_X^\lambda)^{-1}, \tag{4.5.8}$$

其中 $\widehat{\mathcal{C}}$ 和 $\mathcal{C}$ 的定义同定义 4.4.3, $\widehat{\mathcal{C}}$ 对应的 Y 即 $\widehat{Z}$,$\mathcal{C}$ 对应的 Y 即 $Z = \mathbf{R}^N$. 设

$$V = \frac{\mathrm{d}}{\mathrm{d}\lambda}\widetilde{\widehat{\phi}}\Big|_{\lambda=0} = V^{\widehat{\phi}}\frac{\partial}{\partial\widehat{\phi}} = V^{\phi}\frac{\partial}{\partial\phi} + V^{\boldsymbol{s}}\frac{\partial}{\partial\boldsymbol{s}} \tag{4.5.9}$$

表示光滑向量场, 它是 η_Y^λ 的无穷小生成子, 其中 V^ϕ 是 η_Z^λ 的无穷小生成子. 对作用函数 (4.5.6) 作变分, 就是寻求临界截面 ϕ, 使得 $\dfrac{\mathrm{d}}{\mathrm{d}\lambda}S(\widetilde{\phi},\widetilde{x},\widetilde{t})\Big|_{\lambda=0} = 0$, 计算如下:

$$\begin{aligned}
\mathrm{d}S(\widetilde{\phi},\widetilde{x},\widetilde{t}) &= \frac{\mathrm{d}}{\mathrm{d}\lambda}S(\widetilde{\phi},\widetilde{x},\widetilde{t})\Big|_{\lambda=0}\\
&= \int_U \frac{\mathrm{d}}{\mathrm{d}\lambda}\left[\mathfrak{F}(\widetilde{\boldsymbol{s}},\widetilde{\phi},\widetilde{\phi}_t) + \mathfrak{G}(\widetilde{\boldsymbol{s}},\widetilde{\phi},\widetilde{\phi}_x) - \mathfrak{B}(\widetilde{\phi},\widetilde{\boldsymbol{s}},\widetilde{\boldsymbol{s}}_t,\widetilde{\boldsymbol{s}}_x)\right]\Big|_{\lambda=0}\mathrm{d}x\wedge\mathrm{d}t\\
&= \int_U \left[\frac{\partial\mathfrak{F}}{\partial\phi}V^\phi + \frac{\partial\mathfrak{F}}{\partial\phi_t}\left(\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\mathrm{d}\widetilde{\phi}}{\mathrm{d}t}\Big|_{\lambda=0}\right) + \frac{\partial\mathfrak{G}}{\partial\phi}V^\phi + \frac{\partial\mathfrak{G}}{\partial\phi_x}\left(\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\mathrm{d}\widetilde{\phi}}{\mathrm{d}x}\Big|_{\lambda=0}\right)\right.\\
&\quad + \frac{\partial\mathfrak{F}}{\partial\boldsymbol{s}}V^{\boldsymbol{s}} + \frac{\partial\mathfrak{G}}{\partial\boldsymbol{s}}V^{\boldsymbol{s}} - \frac{\partial\mathfrak{B}}{\partial\phi}V^\phi - \frac{\partial\mathfrak{B}}{\partial\boldsymbol{s}}V^{\boldsymbol{s}} - \frac{\partial\mathfrak{B}}{\partial\boldsymbol{s}_t}\left(\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\mathrm{d}\widetilde{\boldsymbol{s}}}{\mathrm{d}t}\Big|_{\lambda=0}\right)\\
&\quad \left.- \frac{\partial\mathfrak{B}}{\partial\boldsymbol{s}_x}\left(\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\mathrm{d}\widetilde{\boldsymbol{s}}}{\mathrm{d}x}\Big|_{\lambda=0}\right)\right]\mathrm{d}x\wedge\mathrm{d}t\\
&= \int_U\left[\frac{\partial\mathfrak{F}}{\partial\phi} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathfrak{F}}{\partial\phi_t}\right) + \frac{\partial\mathfrak{G}}{\partial\phi} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial\mathfrak{G}}{\partial\phi_x}\right) - \frac{\partial\mathfrak{B}}{\partial\phi}\right]V^\phi\\
&\quad + \int_U\left[\frac{\partial\mathfrak{F}}{\partial\boldsymbol{s}} + \frac{\partial\mathfrak{G}}{\partial\boldsymbol{s}} - \frac{\partial\mathfrak{B}}{\partial\boldsymbol{s}} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathfrak{B}}{\partial\boldsymbol{s}_t}\right) + \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial\mathfrak{B}}{\partial\boldsymbol{s}_x}\right)\right]V^{\boldsymbol{s}}\\
&\quad + \int_{\partial U}\left[\left(\frac{\partial\mathfrak{F}}{\partial\phi_t}\mathrm{d}x + \frac{\partial\mathfrak{G}}{\partial\phi_x}\mathrm{d}t\right)V^\phi - \left(\frac{\partial\mathfrak{B}}{\partial\boldsymbol{s}_t}\mathrm{d}x + \frac{\partial\mathfrak{B}}{\partial\boldsymbol{s}_x}\mathrm{d}t\right)V^{\boldsymbol{s}}\right].
\end{aligned} \tag{4.5.10}$$

在上面的变分过程中, 包含对 V 的一阶延拓作用, V 的一阶延拓表示为

$$\mathrm{D}V = \mathrm{D}_t V + \mathrm{D}_x V = \mathrm{D}_t(V^\phi + V^{\boldsymbol{s}}) + \mathrm{D}_x(V^\phi + V^{\boldsymbol{s}}). \tag{4.5.11}$$

假设作用函数 $S(\widehat{\phi})$ 在群 G 作用下是不变的, 即 $\widehat{\mathrm{d}}S = 0$, 那么区域内部和边界积分分别等于零. 因为 $V^{\widehat{\phi}} = V^\phi + V^{\boldsymbol{s}}$ 是任意的, 所以在方向 V^ϕ 和 $V^{\boldsymbol{s}}$ 中得到两个方程 (组):

$$V^\phi: \frac{\partial \mathfrak{F}}{\partial \phi} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathfrak{F}}{\partial \phi_t}\right) + \frac{\partial \mathfrak{G}}{\partial \phi} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial \mathfrak{G}}{\partial \phi_x}\right) - \frac{\partial \mathfrak{B}}{\partial \phi} = 0, \tag{4.5.12}$$

$$V^{\boldsymbol{s}}: \frac{\partial \mathfrak{F}}{\partial \boldsymbol{s}} + \frac{\partial \mathfrak{G}}{\partial \boldsymbol{s}} - \frac{\partial \mathfrak{B}}{\partial \boldsymbol{s}} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathfrak{B}}{\partial \boldsymbol{s}_t}\right) + \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial \mathfrak{B}}{\partial \boldsymbol{s}_x}\right) = 0. \tag{4.5.13}$$

在坐标意义下, 将 $s_1 = t$ 和 $s_2 = x$ 代入 $\mathfrak{B}$, 则方程 (4.5.12) 就是变分 (4.5.1) 对应的 Birkhoff 型 Euler-Lagrange 方程. 因为

$$\mathfrak{B}(x, t, \phi, \boldsymbol{s}, \boldsymbol{s}_t, \boldsymbol{s}_x) = \frac{1}{2}B(s_1, s_2, \phi)(s_1)_t + \frac{1}{2}B(s_1, s_2, \phi)(s_2)_x, \tag{4.5.14}$$

所以沿 $V^{\boldsymbol{s}}$ 方向的方程 (4.5.13) 可以分解到 V^{s_1} 和 V^{s_2} 两个方向. 同时, 由 (4.5.10) 式可给出如下 **Cartan 形式**:

$$\Theta_L = \frac{\partial \mathfrak{F}}{\partial \phi_t}\mathrm{d}\phi \wedge \mathrm{d}x - \frac{\partial \mathfrak{G}}{\partial \phi_x}\mathrm{d}\phi \wedge \mathrm{d}t - \mathfrak{B}\mathrm{d}t \wedge \mathrm{d}x. \tag{4.5.15}$$

对 $S(\phi, x, t)$ 作全变分, 在 V^ϕ 方向得到和方程 (4.5.12) 完全相同的 Euler-Lagrange 方程, 在 V^t 和 V^x 方向产生两个新的方程, 用局部坐标表示分别为

$$V^t: \frac{\partial F_\nu}{\partial t}z_t^\nu + \frac{\partial G_\nu}{\partial t}z_x^\nu - \frac{\partial B}{\partial t} + \frac{\mathrm{d}}{\mathrm{d}t}(B - G_\nu z_x^\nu) + + \frac{\mathrm{d}}{\mathrm{d}x}(G_\nu z_t^\nu) = 0, \tag{4.5.16}$$

$$V^x: \frac{\partial F_\nu}{\partial x}z_t^\nu + \frac{\partial G_\nu}{\partial x}z_x^\nu - \frac{\partial B}{\partial x} + \frac{\mathrm{d}}{\mathrm{d}x}(B - F_\nu z_t^\nu) + + \frac{\mathrm{d}}{\mathrm{d}t}(F_\nu z_x^\nu) = 0. \tag{4.5.17}$$

它们正是方程组 (4.5.13) 分别在 V^{s_1} 和 V^{s_2} 两个方向对应的方程. 如果

$$\frac{\partial F_\nu}{\partial t}z_t^\nu + \frac{\partial G_\nu}{\partial t}z_x^\nu - \frac{\partial B}{\partial t} = 0,$$

则由方程 (4.5.16) 可以得到形式能量守恒律; 如果

$$\frac{\partial F_\nu}{\partial x} z_t^\nu + \frac{\partial G_\nu}{\partial x} z_x^\nu - \frac{\partial B}{\partial x} = 0,$$

同样可以由方程 (4.5.17) 得到形式动量守恒律.

方程 (4.5.16) 和 (4.5.17) 可以整理变为下面的形式:

$$V^t : \left(\frac{\partial G_\nu}{\partial z^\mu} - \frac{\partial G_\mu}{\partial z^\nu}\right) z_x^\nu z_t^\mu - \frac{\partial B}{\partial z^\mu} z_t^\mu - \frac{\partial F_\mu}{\partial t} z_t^\mu - \frac{\partial G_\mu}{\partial x} z_t^\mu = 0, \tag{4.5.18}$$

$$V^x : \left(\frac{\partial F_\nu}{\partial z^\mu} - \frac{\partial F_\mu}{\partial z^\nu}\right) z_t^\nu z_x^\mu - \frac{\partial B}{\partial z^\mu} z_x^\mu - \frac{\partial F_\mu}{\partial t} z_x^\mu - \frac{\partial G_\mu}{\partial x} z_x^\mu = 0. \tag{4.5.19}$$

显然, 方程 (4.5.18) 和 (4.5.19) 是 Euler-Lagrange 方程 (4.5.12) 两边分别乘以 z_t^μ 和 z_x^μ 以及利用 $\left(\frac{\partial F_\nu}{\partial z^\mu} - \frac{\partial F_\mu}{\partial z^\nu}\right)$ 和 $\left(\frac{\partial G_\nu}{\partial z^\mu} - \frac{\partial G_\mu}{\partial z^\nu}\right)$ 的反对称性得到的. 由边界上的积分同样得到 Cartan 形式 (4.5.15).

令

$$\theta_L = \frac{\partial \mathfrak{F}}{\partial \phi_t} \mathrm{d}\phi \wedge \mathrm{d}x - \frac{\partial \mathfrak{G}}{\partial \phi_x} \mathrm{d}\phi \wedge \mathrm{d}t, \quad \theta_{\boldsymbol{s}} = -\mathfrak{B} \mathrm{d}t \wedge \mathrm{d}x, \tag{4.5.20}$$

那么有

$$\Theta_L = \theta_L + \theta_{\boldsymbol{s}}, \quad \Omega_L = \widehat{\mathrm{d}}(\theta_L + \theta_{\boldsymbol{s}}) = \omega_L + \omega_{\boldsymbol{s}}, \quad \omega_{\boldsymbol{s}} = \mathrm{d}^{\boldsymbol{s}}\theta_L + \widehat{\mathrm{d}}\theta_{\boldsymbol{s}}. \tag{4.5.21}$$

记关于 ϕ 的 Euler-Lagrange 算子为 D_{EL}^ϕ, 而关于 $\boldsymbol{x}$ 的 Euler-Lagrange 算子为 $D_{EL}^{\boldsymbol{x}}$, 那么我们得到以下定理:

定理 4.5.2 假设 Pfaff 作用泛函 (4.5.1) 给定的一阶 Lagrange 函数 $L : J^1(Y) \to \Lambda^{n+1}(X)$ 是光滑的, 那么存在唯一的 C^∞ 光滑截面 $D_{EL}L : Y' \to \Lambda^{n+1}(X) \otimes T^*Y$ 以及唯一的微分形式 $\Theta_L \in \Lambda^{n+1}(J^1(Y))$, 使得对于任意的 $V = V^\phi + V^{\boldsymbol{s}}$ 和 X 的任意开子集 U_X ($\overline{U}_X \cap \partial X = \varnothing$), 成立

$$\mathrm{d}S_{\widehat{\phi}} \cdot V = \int_{U_X} D_{EL}L(j^1(\widehat{\phi} \circ \widehat{\phi}_X^{-1})) \cdot V + \int_{\partial U_X} j^1(\widehat{\phi} \circ \widehat{\phi}_X^{-1})^*(i_{j^1(V)}\Theta_L), \tag{4.5.22}$$

且在 U_X 上有

$$D_{EL}L(j^1(\widehat{\phi}\circ\widehat{\phi}_X^{-1}))\cdot V = j^1(\widehat{\phi}\circ\widehat{\phi}_X^{-1})^*(i_{j^1(V)}\Omega_L), \tag{4.5.23}$$

即 Y' 上的 Euler-Lagrange 算子作用在 L 上得到的 $D_{EL}L$ 可用局部坐标表示为

$$D_{EL}L = D_{EL}^{\phi}L + D_{EL}^{\boldsymbol{x}}L,$$

其中

$$\begin{aligned}
&D_{EL}^{\phi}L(j^1(\widehat{\phi}\circ\widehat{\phi}_X^{-1}))\\
&\quad=\left[\frac{\partial L}{\partial y^A}(j^1(\phi\circ\phi_X^{-1})) - \frac{\partial^2 L}{\partial x^\mu \partial v_\mu^A}(j^1(\phi\circ\phi_X^{-1}))\right.\\
&\qquad\left.-\frac{\partial^2 L}{\partial y^B \partial v_\mu^A}(j^1(\phi\circ\phi_X^{-1}))\cdot(\phi\circ\phi_X^{-1})_{,\mu}^B\right]\mathrm{d}y^A\wedge\mathrm{d}^n\boldsymbol{x}\wedge\mathrm{d}t,\\
&D_{EL}^{\boldsymbol{x}}L(j^1(\widehat{\phi}\circ\widehat{\phi}_X^{-1}))\\
&\quad=\left[\frac{\partial \mathfrak{F}}{\partial x^\mu}(j^1(\phi\circ\phi_X^{-1}))+\frac{\partial \mathfrak{B}}{\partial y^B}(j^1(\phi\circ\phi_X^{-1}))\cdot(\phi\circ\phi_X^{-1})_{,\mu}^B\right]\mathrm{d}x^\mu\wedge\mathrm{d}^n\boldsymbol{x}\wedge\mathrm{d}t,
\end{aligned} \tag{4.5.24}$$

这里 $\Omega_L = \mathrm{d}\Theta_L$, Θ_L 即 Cartan 形式, 其坐标表示为

$$\Theta_L = \frac{\partial L}{\partial v_\mu^A}\mathrm{d}y^A\wedge\mathrm{d}^n x_\mu + \left(L - \frac{\partial L}{\partial v_\mu^A}v_\mu^A\right)\mathrm{d}^n\boldsymbol{x}\wedge\mathrm{d}t. \tag{4.5.25}$$

定义 4.5.1 定义集合

$$\widehat{\mathcal{P}} = \{\widehat{\phi}\in\widehat{\mathcal{C}}\mid j^1(\widehat{\phi}\circ\widehat{\phi}_X^{-1})^*(i_W\Omega_L)=0,\ \text{对所有的}\ W\in TJ^1(Y)\}, \tag{4.5.26}$$

对于任意的 $\widehat{\phi}\in\widehat{\mathcal{P}}$, 定义集合

$$\begin{aligned}\widehat{\mathcal{F}} = \{V\in T_{\widehat{\phi}}\widehat{\mathcal{C}}\mid & j^1(\widehat{\phi}\circ\widehat{\phi}_X^{-1})^* L_{j^1(V)}(i_W\Omega_L)=0,\\ &\text{对所有的}\ W\in TJ^1(Y)\}.\end{aligned} \tag{4.5.27}$$

定理 4.5.3 如果 $\widehat{\phi}\in\widehat{\mathcal{C}}$, 则对任意的 $V, W\in\widehat{\mathcal{F}}$, 成立

$$\int_{\partial U} j^1(\widehat{\phi}\circ\widehat{\phi}_X^{-1})^*(i_{j^1(V)}(i_{j^1(W)}\Omega_L)) = 0. \tag{4.5.28}$$

定理证明类似于非奇异 Lagrange 变分问题的同一结论, 可参看文献 [MW01, MPS98].

§4.6 空气中带磨擦的弦振动方程的变分积分子

在这一节中, 我们以空气中带磨擦的弦振动方程为例, 讨论有外力作用的变分问题的几何结构及其离散形式. 考虑方程

$$u_{tt} - u_{xx} + u + 2u_t = 0, \tag{4.6.1}$$

其中有耗散项 u_t. 引进势函数 p 和 q, 原方程等价于

$$u_t = p, \quad u_x = q, \quad p_t - q_x + u + 2p = 0, \tag{4.6.2}$$

相应的 Birkhoff 函数是

$$\begin{aligned}
&\boldsymbol{F} = \left(-\frac{1}{2}\mathrm{e}^{2t}p, \frac{1}{2}\mathrm{e}^{2t}u, 0\right)^{\mathrm{T}}, \quad \boldsymbol{G} = \left(\frac{1}{2}\mathrm{e}^{2t}q, 0, -\frac{1}{2}\mathrm{e}^{2t}u\right)^{\mathrm{T}}, \\
&B = -\frac{1}{2}\mathrm{e}^{2t}(u^2 + p^2 - q^2 + 2up).
\end{aligned} \tag{4.6.3}$$

令 $X = \{(x,t)\}$, $Z = \{\boldsymbol{z} = (u,p,q)\}$, $Y = \{(u,p,q,s_1 = t, s_2 = x)\}$, 则对应方程组 (4.6.2) 的作用泛函定义为

$$\begin{aligned}
S &= \int_X L(j^1(\phi))\mathrm{d}x \wedge \mathrm{d}s_1 \\
&= \int_X [F_\nu(u,p,q,s_1,s_2)z_t^\nu + G_\nu(u,p,q,s_1,s_2)z_x^\nu \\
&\quad -B(u,p,q,s_1,s_2)]\mathrm{d}x \wedge \mathrm{d}t \\
&= \int_X \left[\frac{1}{2}\mathrm{e}^{2s_1}(-pu_t + up_t + qu_x - uq_x) + \frac{1}{4}\mathrm{e}^{2s_1}(u^2 + p^2 - q^2 + 2up)(s_1)_t \right. \\
&\quad \left. + \frac{1}{4}\mathrm{e}^{2s_1}(u^2 + p^2 - q^2 + 2up)(s_2)_x\right]\mathrm{d}x \wedge \mathrm{d}t.
\end{aligned} \tag{4.6.4}$$

令 $V = V^{\boldsymbol{s}} + V^{\boldsymbol{z}} = V^{\boldsymbol{s}} + V^u + V^p + V^q$. 由于 $s_1 = t$, $s_2 = x$, 所以 V 的一阶延拓定义如下:

$$j^1(V) = V + \mathrm{D}_t V^{\boldsymbol{z}} + \mathrm{D}_x V^{\boldsymbol{z}} + \mathrm{D}_t V^{s_1} + \mathrm{D}_x V^{s_2}. \tag{4.6.5}$$

再利用 Green 公式, 得到

$$\mathrm{d}S = \int_X I_1 + \int_{\partial X} I_2, \tag{4.6.6}$$

其中

$$
\begin{aligned}
I_1 = \frac{1}{2}\mathrm{e}^{2s_1}\Big[&(2p_t - 2q_x + u(s_1)_t + u(s_2)_x + 3p(s_1)_t + p(s_2)_x)V^u \\
&+(-2u_t + p(s_1)_t + p(s_2)_x)V^p + (2u_x - q(s_1)_t - q(s_2)_x)V^q \\
&+\Big(- 3pu_t + up_t + 2qu_x - 2uq_x - uu_t - pp_t + qq_t \\
&-\frac{1}{2}(u^2 + p^2 - q^2 + 2up)((s_1)_t - (s_2)_x)\Big)V^{s_1} \\
&-(uu_x + pp_x - qq_x + up_x + pu_x)V^{s_2}\Big]\mathrm{d}x \wedge \mathrm{d}t,
\end{aligned} \tag{4.6.7}
$$

$$
\begin{aligned}
I_2 = &\left[\frac{1}{4}\mathrm{e}^{2t}(-2pV^u + 2uV^p + (u^2 + p^2 - q^2 + 2up)V^{s_1})\right]\mathrm{d}x \\
&- \left[\frac{1}{4}\mathrm{e}^{2t}(uV^q - qV^u + (u^2 + p^2 - q^2 + 2up)V^{s_2})\right]\mathrm{d}t.
\end{aligned} \tag{4.6.8}
$$

将 $s_1 = t$ 和 $s_2 = x$ 代入 I_2, 定义 Cartan 形式:

$$
\begin{aligned}
\Theta_L = &\frac{1}{2}\mathrm{e}^{2t}(-p\mathrm{d}u \wedge \mathrm{d}x + u\mathrm{d}p \wedge \mathrm{d}x + q\mathrm{d}u \wedge \mathrm{d}t - u\mathrm{d}q \wedge \mathrm{d}t) \\
&+\frac{1}{2}\mathrm{e}^{2t}(u^2 + p^2 - q^2 + 2up)\mathrm{d}t \wedge \mathrm{d}x.
\end{aligned} \tag{4.6.9}
$$

将 $s_1 = t$ 和 s_2 代入 I_1 中沿 V^u, V^p 和 V^q 方向的变量, 得到 Euler-Lagrange 方程组等价于方程组 (4.6.2), 沿着 V^{s_1} 和 V^{s_2} 得到形式能量和动量公式.

给定离散 $1+1$ 维时–空微分流形 X, 令 $X = \{(x_i, t_j)\}$. 为了简便, 以下用 (i,j) 记 (x_i, t_j), 对于 X 上的任意函数 f, 记 $f_{i,j} = f(x_i, t_j)$. 取纤维丛 Y 为 $X \times \mathcal{R}$, 其中 $\mathcal{R}$ 为三维光滑微分流形, 则 (i,j) 处 Y 的元素表示为 $(\phi_{i,j}, \psi_{i,j}, \xi_{i,j})$. X 中的一个矩形是一个有序四元组:

$$
\square_{i,j} = ((i,j), (i,j+1), (i+1,j+1), (i+1,j)). \tag{4.6.10}
$$

$\square_{i,j}$ 中的分量称为矩形 $\square_{i,j}$ 的顶点. 如果点 (i,j) 是一个矩形的顶点, 则称这个点和该矩形接触. 设 $U \subseteq X$, 如果和点 (i,j) 接触的四个矩形都包含于 U 内部, 则称 (i,j) 是 U 的内点. 用 $\overline{U}$ 表示和 U 的所有内点接触的矩形的并. U 的边界点就是 $\overline{U}$ 中除去 U 的内点的那些点. 用 ∂U 表示 U 的边界点的集合.

Y 的离散一阶 Jet 丛定义为

$$J^1(Y)=(i,j;\phi_{i,j},\psi_{i,j},\xi_{i,j},\phi_{i,j+1},\psi_{i,j+1},\xi_{i,j+1},\phi_{i+1,j+1},\\ \psi_{i+1,j+1},\xi_{i+1,j+1},\phi_{i+1,j},\psi_{i+1,j},\xi_{i+1,j}), \tag{4.6.11}$$

Y 的截面 $\boldsymbol{\varphi}_{i,j}=(u_{i,j},p_{i,j},q_{i,j})$ 的一阶 Jet 为

$$j^1(\boldsymbol{\varphi})(i,j)=(i,j;u_{i,j},p_{i,j},q_{i,j},u_{i,j+1},p_{i,j+1},q_{i,j+1},u_{i+1,j+1},\\ p_{i+1,j+1},q_{i+1,j+1},u_{i+1,j},p_{i+1,j},q_{i+1,j}). \tag{4.6.12}$$

设 h 和 k 分别为时间和空间步长. 利用双线性插值, 得到 φ 在矩形 $\square_{i,j}$ 上的一个插值函数 $\widetilde{\boldsymbol{\varphi}}$:

$$\begin{aligned}\widetilde{\boldsymbol{\varphi}}(x,t)=&\frac{\boldsymbol{\varphi}_{i,j}}{hk}(x-x_{i+1,j})(t-t_{i,j+1})-\frac{\boldsymbol{\varphi}_{i,j+1}}{hk}(x-x_{i+1,j+1})(t-t_{i,j})\\ &+\frac{\boldsymbol{\varphi}_{i+1,j+1}}{hk}(x-x_{i,j+1})(t-t_{i+1,j})\\ &-\frac{\boldsymbol{\varphi}_{i+1,j}}{hk}(x-x_{i,j})(t-t_{i+1,j+1}).\end{aligned} \tag{4.6.13}$$

直接计算得到

$$\begin{aligned}\frac{\partial\widetilde{\boldsymbol{\varphi}}}{\partial x}=&\frac{\boldsymbol{\varphi}_{i,j}}{hk}(t-t_{i,j+1})-\frac{\boldsymbol{\varphi}_{i,j+1}}{hk}(t-t_{i,j})\\ &+\frac{\boldsymbol{\varphi}_{i+1,j+1}}{hk}(t-t_{i+1,j})-\frac{\boldsymbol{\varphi}_{i+1,j}}{hk}(t-t_{i+1,j+1}),\\ \frac{\partial\widetilde{\boldsymbol{\varphi}}}{\partial t}=&\frac{\boldsymbol{\varphi}_{i,j}}{hk}(x-x_{i+1,j})-\frac{\boldsymbol{\varphi}_{i,j+1}}{hk}(x-x_{i+1,j+1})\\ &+\frac{\boldsymbol{\varphi}_{i+1,j+1}}{hk}(x-x_{i,j+1})-\frac{\boldsymbol{\varphi}_{i+1,j}}{hk}(x-x_{i,j}).\end{aligned} \tag{4.6.14}$$

令

$$\widetilde{x}=(x_{i,j}+x_{i,j+1}+x_{i+1,j+1}+x_{i+1,j})/4,$$
$$\widetilde{t}=(t_{i,j}+t_{i,j+1}+t_{i+1,j+1}+t_{i+1,j})/4,$$

代入 (4.6.13) 式和 (4.6.14) 式, 得到

$$\begin{aligned}&\widetilde{\boldsymbol{\varphi}}\square_{i,j}=\widetilde{\boldsymbol{\varphi}}(\widetilde{x},\widetilde{t})=\frac{\boldsymbol{\varphi}_{i,j}+\boldsymbol{\varphi}_{i,j+1}+\boldsymbol{\varphi}_{i+1,j+1}+\boldsymbol{\varphi}_{i+1,j}}{4},\\ &\widetilde{\boldsymbol{\varphi}}_x\square_{i,j}=\frac{\partial\widetilde{\boldsymbol{\varphi}}}{\partial x}(\widetilde{x},\widetilde{t})=\frac{1}{2}\left(\frac{\boldsymbol{\varphi}_{i+1,j}-\boldsymbol{\varphi}_{i,j}}{k}+\frac{\boldsymbol{\varphi}_{i+1,j+1}-\boldsymbol{\varphi}_{i,j+1}}{k}\right),\\ &\widetilde{\boldsymbol{\varphi}}_t\square_{i,j}=\frac{\partial\widetilde{\boldsymbol{\varphi}}}{\partial t}(\widetilde{x},\widetilde{t})=\frac{1}{2}\left(\frac{\boldsymbol{\varphi}_{i,j+1}-\boldsymbol{\varphi}_{i,j}}{h}+\frac{\boldsymbol{\varphi}_{i+1,j+1}-\boldsymbol{\varphi}_{i+1,j}}{h}\right).\end{aligned} \tag{4.6.15}$$

以上三个式子分别表示 φ, φ_x 和 φ_t 在以 (i,j) 为第一个顶点的矩形 $\square_{i,j}$ 上的离散. 那么对应泛函 (4.6.4) 的离散的 Lagrange 函数是

$$L=\frac{1}{2}\mathrm{e}^{\widetilde{t}}(-\widetilde{p}\,\widetilde{u}_t+\widetilde{u}\,\widetilde{p}_t+\widetilde{q}\,\widetilde{u}_x-\widetilde{u}\widetilde{q}_x+\widetilde{u}^2+\widetilde{p}^2-\widetilde{q}^2+2\widetilde{p}\,\widetilde{u}). \tag{4.6.16}$$

上面的 L 并不依赖于具体的矩形, 具体到矩形 $\square_{i,j}$ 时, 表示为 $L(\square_{i,j})$, 那么离散泛函定义为

$$S=\sum_{(i,j)\in U}L(\square_{i,j}). \tag{4.6.17}$$

给定 Y 上的向量场 V, 它的流为 F_V^{ε}, 那么存在一个单参数截面族

$$(F_V^{\varepsilon}(\varphi))(i,j)=F_V^{\varepsilon}(\varphi_{i,j}). \tag{4.6.18}$$

离散的变分原理就是求 S 的临界点, 使得

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}S(F_V^{\varepsilon}(\varphi))\Big|_{\varepsilon=0}=0,\quad \forall V, \tag{4.6.19}$$

即对 (4.6.17) 式关于 $\varphi_{i,j}$ 求变分. 用 $\delta\varphi_{i,j}$ 表示 $\varphi_{i,j}$ 的变分, 又因为点 (i,j) 和四个矩形接触, 所以沿 $\varphi_{i,j}$ 有四项相加, 即

$$\begin{aligned}\delta S=\sum_{(i,j)\text{是 } U \text{ 的内点}}&\frac{\partial}{\partial\varphi_{i,j}}[L(\square_{i,j})+L(\square_{i,j-1})+L(\square_{i-1,j})\\&+L(\square_{i-1,j-1})]\delta\varphi_{i,j}.\end{aligned} \tag{4.6.20}$$

于是得到离散的 Euler-Lagrange 方程

$$\begin{aligned}&\frac{\partial}{\partial u_{i,j}}[L(\square_{i,j})+L(\square_{i,j-1})+L(\square_{i-1,j})+L(\square_{i-1,j-1})]=0,\\&\frac{\partial}{\partial p_{i,j}}[L(\square_{i,j})+L(\square_{i,j-1})+L(\square_{i-1,j})+L(\square_{i-1,j-1})]=0,\\&\frac{\partial}{\partial q_{i,j}}[L(\square_{i,j})+L(\square_{i,j-1})+L(\square_{i-1,j})+L(\square_{i-1,j-1})]=0,\end{aligned} \tag{4.6.21}$$

我们容易推导出 L 关于每个矩形 $\square_{i,j}$ 的第一和第二个顶点求偏导数一定相等, 而关于第三和第四个顶点求偏导数也一定相等, 分别记为

$\dfrac{\partial L}{\partial \varphi_{1,2}}$ 和 $\dfrac{\partial L}{\partial \varphi_{3,4}}$, 所以

$$
\begin{aligned}
\frac{\partial L}{\partial u_{1,2}} &= \left[\frac{1}{2}\mathrm{e}^{2\widetilde{t}}\left(\frac{\widetilde{p}}{2h}+\frac{\widetilde{p}_t}{4}-\frac{\widetilde{q}}{2k}-\frac{\widetilde{q}_x}{2k}+\frac{\widetilde{u}}{2}+\frac{\widetilde{p}}{2}\right)\right],\\
\frac{\partial L}{\partial u_{3,4}} &= \left[\frac{1}{2}\mathrm{e}^{2\widetilde{t}}\left(-\frac{\widetilde{p}}{2h}+\frac{\widetilde{p}_t}{4}+\frac{\widetilde{q}}{2k}-\frac{\widetilde{q}_x}{2k}+\frac{\widetilde{u}}{2}+\frac{\widetilde{p}}{2}\right)\right],\\
\frac{\partial L}{\partial p_{1,2}} &= \left[\frac{1}{2}\mathrm{e}^{2\widetilde{t}}\left(-\frac{\widetilde{u}_t}{4}-\frac{\widetilde{u}}{2h}+\frac{\widetilde{p}}{2}+\frac{\widetilde{u}}{2}\right)\right],\\
\frac{\partial L}{\partial p_{3,4}} &= \left[\frac{1}{2}\mathrm{e}^{2\widetilde{t}}\left(-\frac{\widetilde{u}_t}{4}+\frac{\widetilde{u}}{2h}+\frac{\widetilde{p}}{2}+\frac{\widetilde{u}}{2}\right)\right],\\
\frac{\partial L}{\partial q_{1,2}} &= \left[\frac{1}{2}\mathrm{e}^{2\widetilde{t}}\left(\frac{\widetilde{u}_x}{4}+\frac{\widetilde{u}}{2k}-\frac{\widetilde{q}}{2}\right)\right],\\
\frac{\partial L}{\partial q_{3,4}} &= \left[\frac{1}{2}\mathrm{e}^{2\widetilde{t}}\left(\frac{\widetilde{u}_x}{4}-\frac{\widetilde{u}}{2k}-\frac{\widetilde{q}}{2}\right)\right].
\end{aligned}
\tag{4.6.22}
$$

根据离散 Euler-Lagrange 方程, 我们可以得到

$$
\begin{aligned}
&\frac{\partial}{\partial u_{1,2}}L(\square_{i,j})+\frac{\partial}{\partial u_{1,2}}L(\square_{i,j-1})+\frac{\partial}{\partial u_{3,4}}L(\square_{i-1,j-1})+\frac{\partial}{\partial u_{3,4}}L(\square_{i-1,j})=0,\\
&\frac{\partial}{\partial p_{1,2}}L(\square_{i,j})+\frac{\partial}{\partial p_{1,2}}L(\square_{i,j-1})+\frac{\partial}{\partial p_{3,4}}L(\square_{i-1,j-1})+\frac{\partial}{\partial p_{3,4}}L(\square_{i-1,j})=0,\\
&\frac{\partial}{\partial q_{1,2}}L(\square_{i,j})+\frac{\partial}{\partial q_{1,2}}L(\square_{i,j-1})+\frac{\partial}{\partial q_{3,4}}L(\square_{i-1,j-1})+\frac{\partial}{\partial q_{3,4}}L(\square_{i-1,j})=0.
\end{aligned}
\tag{4.6.23}
$$

将离散格式 (4.6.15) 和 (4.6.22) 式代入 (4.6.23) 式, 得到方程组 (4.6.2) 的一个九点积分子. 因为它是由离散变分得到的, 所以称之为**变分积分子**. 消去 p 和 q 就得到振动方程 (4.6.1) 的一个九点差分格式.

参 考 文 献

[MW01] Marsden J E, West M. Discrete mechanics and variational integrators. *Acta Numerica*, 2001, 10: 357–514.

[EELMS] Echeverría-Enríquez A, López C, Marín-Solano J, Muñoz-Lecanda M C, Román-Roy N. Lagrangian-Hamiltonian unified formalism for field theory. *J Math Phys*, 2004, 45: 360–380.

[KMOW00] Kane C, Marsden J E, Ortiz M, West M. Variational integrators and the Newmark algorithm for conservative and dissipative mechanical systems. *J Numer Math Eng*, 2000, 49: 1295–1325.

[Arn63] Arnold V I. Proof of a theorem of A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian. *Russ Math Surv*, 1963, 18(5): 9–36.

[Mos62] Moser J. On invariant curves of area-preserving mappings of an annulus. *Nachr Akad Wiss Gottingen, II, Math Phys*, 1962: 1–20.

[San83b] Santilli R M. *Foundations of Theoretical Mechanics II*. 2nd edition. New York: Springer-Verlag, 1983.

[CGW03] Chen J B, Guo H Y, Wu K. Total variation in Hamiltonian formalism and symplectic-energy integrators. *J Math Phys*, 2003, 44: 1688–1702.

[FQ91] Feng K, Qin M Z. Hamiltonian algorithms for Hamiltonian dynamical systems. *Progr Natur Sci*, 1991, 1(2): 105–116.

[MPS98] Marsden J E, Patrick G P, Shkoller S. Multisymplectic geometry, variational integrators, and nonlinear PDEs. *Commun Math Phys*, 1998, 199: 351–395.

[Ves88] Veselov A P. Integrable discrete-time systems and difference operators. *Funkts Anal Prilozhen*, 1988, 22: 1–33.

[Ves91] Veselov A P. Integrable maps. *Russ Math Surv*, 1991, 46: 1–51.

[WM97] Wendlandt J, Marsden J E. Mechanical integrators derived from a discrete variational principle. *Physica D*, 1997, 106: 223–246.

第五章　有限维 Birkhoff 系统的辛结构和辛格式

Hamilton 系统通常表示一种没有受外力作用的系统. 对于 Hamilton 系统的辛结构, 我们已经有了充分的认识和理解. 对于带有耗散项的方程或受外力作用的系统, 它们是否也具备某种几何结构? 这种结构又是什么样的? 它和通常的 Hamilton 系统的标准辛结构有什么不同? 本章就这几个问题进行讨论.

将最一般的辛结构下的 Hamilton 系统直接推广可以得到一个带耗散项的系统, 这正是文献 [San83b] 介绍的 Birkhoff 系统. 我们将从变分反问题的角度出发, 讨论最一般的协变一阶常微分方程的对称性. 这种对称性是直观的, 又是从本质上表示出 Hamilton 系统的特殊性和 Birkhoff 系统的一般性.

§5.1　有限维 Birkhoff 方程及其几何结构

5.1.1　有限维 Birkhoff 方程

对于有限维 Hamilton 系统, 我们做如下的推广:

$$\left(\frac{\partial F_j}{\partial z_i}-\frac{\partial F_i}{\partial z_j}\right)\frac{\mathrm{d}z_j}{\mathrm{d}t}-\left(\frac{\partial B(\boldsymbol{z},t)}{\partial z_i}+\frac{\partial F_i(\boldsymbol{z},t)}{\partial t}\right)=0, \tag{5.1.1}$$

其中 $\boldsymbol{z}=(z_1,\cdots,z_{2n})^{\mathrm{T}}\in\mathbf{R}^{2n}$. 为了简便, 我们依然沿用爱因斯坦记号习惯.

出于文献 [San83b] 中提到的历史原因, 我们称系统 (5.1.1) 为 **Birkhoff 方程**或 **Birkhoff 系统**. 函数 $B(\boldsymbol{z},t)$ 和 $F_i(\boldsymbol{z},t)$ 统称为 **Birkhoff 函数**, 特别称 $B(\boldsymbol{z},t)$ 为 **Birkhoff 量**. 一个 Newton 系统通过一个 Birkhoff 方程描述, 称之为 **Birkhoff 表示**.

定义 5.1.1 当 Birkhoff 函数 F_i 和 B 都不显含时间变量 t 时, 称 Birkhoff 方程 (5.1.1) 为**自治**的, 此时方程 (5.1.1) 有如下形式:

$$K_{ij}(\boldsymbol{z})\frac{\mathrm{d}z_j}{\mathrm{d}t} - \frac{\partial B(\boldsymbol{z})}{\partial z_i} = 0; \tag{5.1.2}$$

当 Birkhoff 函数 F_i 不显含 t 而 B 显含 t 时, 称 Birkhoff 方程 (5.1.1) 为**半自治**的, 此时方程 (5.1.1) 有如下形式:

$$K_{ij}(\boldsymbol{z})\frac{\mathrm{d}z_j}{\mathrm{d}t} - \frac{\partial B(\boldsymbol{z},t)}{\partial z_i} = 0; \tag{5.1.3}$$

当 Birkhoff 函数 F_i 和 B 都显含时间 t 时, 称 Birkhoff 方程 (5.1.1) 为**非自治**的, 此时它有如下最一般的形式:

$$K_{ij}(\boldsymbol{z},t)\frac{\mathrm{d}z_j}{\mathrm{d}t} - \frac{\partial B(\boldsymbol{z},t)}{\partial z_i} - \frac{\partial F_i(\boldsymbol{z},t)}{\partial t} = 0. \tag{5.1.4}$$

这里无论是自治、半自治或非自治的情形, 都有

$$K_{ij} = \frac{\partial F_j}{\partial z_i} - \frac{\partial F_i}{\partial z_j}. \tag{5.1.5}$$

若 $\boldsymbol{K}(\boldsymbol{z},t) = (K_{ij}(\boldsymbol{z},t))$ 的行列式在某一区域 $\widetilde{\Re}$ 上不等于零, 即

$$\det(\boldsymbol{K})(\widetilde{\Re}) \neq 0, \tag{5.1.6}$$

则称其在该区域是**正则**的; 反之, 称为**退化**的.

事实上, 自治和半自治的 Birkhoff 方程就是通常所说的一般辛结构下的自治和非自治的 Hamilton 方程. 我们后面的讨论针对最一般的 Birkhoff 方程, 即非自治的情形. 所以说, Birkhoff 方程和 Hamilton 方程的不同之处在于, 除了有一个显含时间 t 的结构 $\boldsymbol{K}(\boldsymbol{z},t)$ 外, 它还多出了一个耗散项 $\dfrac{\partial F_i}{\partial t}$.

5.1.2 Birkhoff 结构和 Birkhoff 辛结构

我们考虑任意一个正则的一阶系统

$$N_i(\boldsymbol{z}) \triangleq K_{ij}(\boldsymbol{z},t)\frac{\mathrm{d}z_j}{\mathrm{d}t} + D_i(\boldsymbol{z},t) = 0, \quad i = 1,\cdots,2n, \tag{5.1.7}$$

其中 $D_i(\boldsymbol{z},t)(i=1,\cdots,2n)$ 是标量函数.

在第三章中, 我们从变分反问题的角度出发, 给出了方程对称性的定义. 所以在这里我们可以直接给出方程 (5.1.7) 的对称条件.

定理 5.1.1 方程 (5.1.7) 是对称的当且仅当下列条件满足:

$$\begin{aligned}&K_{ij}+K_{ji}=0,\\&\frac{\partial K_{ij}}{\partial z_k}+\frac{\partial K_{jk}}{\partial z_i}+\frac{\partial K_{ki}}{\partial z_j}=0,\quad i,j,k=1,\cdots,2n.\\&\frac{\partial K_{ij}}{\partial t}=\frac{\partial D_i}{\partial z_j}-\frac{\partial D_j}{\partial z_i},\end{aligned}\tag{5.1.8}$$

条件 (5.1.8) 称为方程 (5.1.7) 的**变分对称条件**.

Hamilton 方程满足变分对称条件 (5.1.8) 的特殊情形, 即满足

$$\frac{\partial D_i}{\partial z_j}=\frac{\partial D_j}{\partial z_i}.$$

在第四章中, 我们从 Pfaff 作用泛函的变分出发讨论了受外力作用的系统的几何性质. 在这一章中, 我们将从微分方程出发, 通过讨论对称条件的代数和几何性质, 将 Hamilton 方程的代数和几何性质推广到更广泛的 Birkhoff 方程框架, 这样自然地将不受外力作用的 Hamilton 系统推广到受外力作用的 Birkhoff 系统.

方程 (5.1.7) 是一个变分问题的 Euler-Lagrange 方程的充分必要条件是它满足可积条件, 即我们上面给出的变分对称条件 (5.1.8). 我们在一个星形区域 $\widetilde{\Re}^*\in\mathbf{R}^{2n}\times\mathbf{R}$ 上进行下面的讨论. 经典的 Poisson 括号是 Newton 力学中满足 Lie 等式 (即反对称性) 和 Jacobi 恒等式的双线性形式的最简单例子. 条件 (5.1.8) 意味着 $\boldsymbol{K}$ 是正则的, 因此不妨假设 $(K^{ij})_{2n\times 2n}=(K_{ij})^{-1}_{2n\times 2n}$, 进而可以给出一个更一般的括号:

$$\{A,B\}=\frac{\partial A}{\partial z^i}K^{ij}\frac{\partial B}{\partial z^j},\quad \det(\boldsymbol{K})(\widetilde{\Re}^*)\neq 0.\tag{5.1.9}$$

这个括号满足反对称性和 Jacobi 恒等式的可积条件是

$$\begin{aligned}&K^{ij}+K^{ji}=0,\\&K^{mn}\frac{\partial K^{ij}}{\partial z^n}+K^{in}\frac{\partial K^{jm}}{\partial z^n}+K^{jn}\frac{\partial K^{mi}}{\partial z^n}=0.\end{aligned}\tag{5.1.10}$$

也就是说, 我们前面给出的方程 (5.1.7) 的变分对称条件定义了一个广义的 Poisson 括号. 下面我们也简单介绍一下它在几何方面的性质.

首先我们考虑自治和半自治的情形. 由 $K_{ij} = K_{ij}(\boldsymbol{z})$ 定义一个 2-形式, 用局部坐标表示为

$$\begin{aligned} \Omega &= \sum_{i,j=1}^{2n} K_{ij}(\boldsymbol{z}) \mathrm{d}z_i \wedge \mathrm{d}z_j, \\ K_{ij} &= -K_{ji}, \quad i,j = 1,\cdots,2n. \end{aligned} \tag{5.1.11}$$

显然, $\mathrm{d}\Omega = 0$ 即 Ω 是一个闭辛形式的可积条件就是变分对称条件 (5.1.8) 中的前面两个等式. 那么 2-形式 (5.1.11) 有下面的结构:

$$\Omega = \mathrm{d}(F_i \mathrm{d}z_i). \tag{5.1.12}$$

闭 2-形式 (5.1.12) 包括了自治和半自治的 Birkhoff 系统的几何结构.

关于非自治的情形, $K_{ij} = K_{ij}(\boldsymbol{z}, t)$, 要特别考虑到变分对称条件 (5.1.8) 中的第三个等式. 前面我们讨论的是 $\mathbf{R}^{2n}$ 的余切丛 $T^*\mathbf{R}^{2n}$ 上的广义辛几何, 对于非自治 (即 $K_{ij} = K_{ij}(\boldsymbol{z}, t)$) 的 Birkhoff 系统, 包括两个几何结构: 一是微分流形 $T^*\mathbf{R}^{2n}$ 上的几何结构 Ω; 二是 $\mathbf{R} \times T^*\mathbf{R}^{2n}$ 上的几何结构 $\widehat{\Omega}$. Ω 还和前面定义的一样, 不过它现在显含参数 t. 我们再看增广空间上的结构 $\widehat{\Omega}$, 它的局部坐标为 $(\widetilde{z}_0, \widetilde{z}_1, \cdots, \widetilde{z}_{2n})$, 其中 $\widetilde{z}_0 = t$. 那么, 在 $2n+1$ 维的局部空间上, 仍旧存在一个闭 2-形式

$$\widehat{\Omega} = \sum_{i,j=0}^{2n} \widehat{K}_{ij} \mathrm{d}\widetilde{z}_i \wedge \mathrm{d}\widetilde{z}_j = \Omega + 2D_i \mathrm{d}z_i \wedge \mathrm{d}t, \tag{5.1.13}$$

这里

$$\widehat{\boldsymbol{K}} = \begin{pmatrix} 0 & -\boldsymbol{D}^{\mathrm{T}} \\ \boldsymbol{D} & \boldsymbol{K} \end{pmatrix}, \tag{5.1.14}$$

$\boldsymbol{D} = (D_1, \cdots, D_{2n})^{\mathrm{T}}$. 2-形式 (5.1.13) 可积, 即

$$\widehat{\Omega} = \mathrm{d}(\widetilde{F}_i \mathrm{d}\widetilde{z}_i), \quad \widetilde{F}_i = \begin{cases} -B, \\ F_i, \end{cases} \quad i = 1, \cdots, 2n, \tag{5.1.15}$$

可见它的可积条件即非自治 Birkhoff 系统的变分对称条件.

定义 5.1.2 Ω 称为 **Birkhoff 辛结构**, $\widehat{\Omega}$ 称为 **Birkhoff 结构**.

显然函数 $B(\boldsymbol{z},t)$ 是一个势函数, 因为

$$-\frac{\partial B}{\partial z_i}=D_i+\frac{\partial F_i}{\partial t} \tag{5.1.16}$$

满足可积条件

$$\frac{\partial}{\partial z_j}\left(D_i+\frac{\partial F_i}{\partial t}\right)=\frac{\partial}{\partial z_i}\left(D_j+\frac{\partial F_j}{\partial t}\right). \tag{5.1.17}$$

Birkhoff 结构可以看做 Birkhoff 辛结构加上约束条件 (5.1.16). 以上讨论可以用一个简单的命题总结.

命题 5.1.1 (Birkhoff 系统的对称性) 一个协变的非自治一阶方程在微分流形 $\mathbf{R}\times T^*\mathbf{R}^{2n}$ 的某一星形区域 $\widetilde{\Re}^*$ 上为对称的充分必要条件是, 它具有 Birkhoff 形式, 即具有以下形式:

$$K_{ij}(\boldsymbol{z},t)\frac{\mathrm{d}z_j}{\mathrm{d}t}+D_i(\boldsymbol{z},t)=\left(\frac{\partial F_j}{\partial z_i}-\frac{\partial F_i}{\partial z_j}\right)\frac{\mathrm{d}z_j}{\mathrm{d}t}-\left(\nabla B(\boldsymbol{z},t)+\frac{\partial \boldsymbol{F}(\boldsymbol{z},t)}{\partial t}\right). \tag{5.1.18}$$

Birkhoff 函数 F_i 和 B 可通过简单的积分运算得到:

$$F_i=\frac{1}{2}\int_0^1 z_jK_{ji}(\lambda\boldsymbol{z},t)\mathrm{d}\lambda, \tag{5.1.19}$$

$$B=-\int_0^1 z_i\left(D_i+\frac{\partial F_i}{\partial t}\right)(\lambda\boldsymbol{z},t)\,\mathrm{d}\lambda. \tag{5.1.20}$$

Birkhoff 方程可以看做 Pfaff 作用泛函变分问题对应的 Euler-Lagrange 方程[San83a,San83b,Arn89]. 而且我们知道, 对于一阶的 Pfaff 作用泛函变分问题, 不存在非平凡的 Legendre 变换可以将受外力作用的一阶 Euler-Lagrange 方程变到不受外力作用的 Hamilton 方程. 我们在这里不详细讨论这一点.

根据系统 (5.1.18) 是对称的, 我们容易证明它的结构 $\widehat{\Omega}$ 满足

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\Omega}=\frac{\mathrm{d}}{\mathrm{d}t}(\widehat{K}_{ij}\mathrm{d}\widetilde{z}_i\wedge\mathrm{d}\widetilde{z}_j)=0. \tag{5.1.21}$$

记方程组 (5.1.18) 在增广空间上的解流为 $(\widehat{\boldsymbol{z}},\widehat{t})$, 那么

$$\begin{aligned}&K_{ij}(\widehat{\boldsymbol{z}},\widehat{t})\mathrm{d}\widehat{z}_i\wedge\mathrm{d}\widehat{z}_j+2D_i(\widehat{\boldsymbol{z}},\widehat{t})\mathrm{d}\widehat{t}\wedge\mathrm{d}\widehat{z}_i\\&=K_{ij}(\boldsymbol{z},t)\mathrm{d}z_i\wedge\mathrm{d}z_j+2D_i(\boldsymbol{z},t)\mathrm{d}t\wedge\mathrm{d}z_i,\end{aligned} \tag{5.1.22}$$

从而由 $\mathrm{d}z_i \wedge \mathrm{d}z_j$ 的系数推出辛流公式

$$K_{ij}(\widehat{\boldsymbol{z}},\widehat{t})\mathrm{d}\widehat{z}_i \wedge \mathrm{d}\widehat{z}_j = K_{ij}(\boldsymbol{z},t)\mathrm{d}z_i \wedge \mathrm{d}z_j, \tag{5.1.23}$$

或者表示为

$$\frac{\partial \widehat{\boldsymbol{z}}}{\partial \boldsymbol{z}}^{\mathrm{T}} \boldsymbol{K}(\widehat{\boldsymbol{z}},\widehat{t})\frac{\partial \widehat{\boldsymbol{z}}}{\partial \boldsymbol{z}} = \boldsymbol{K}(\boldsymbol{z},t); \tag{5.1.24}$$

从 $\mathrm{d}t \wedge \mathrm{d}z_i$ 的系数推出约束条件

$$\frac{\partial \widehat{\boldsymbol{z}}}{\partial t}^{\mathrm{T}} \boldsymbol{K}(\widehat{\boldsymbol{z}},\widehat{t}) + \boldsymbol{D}(\widehat{\boldsymbol{z}},\widehat{t})^{\mathrm{T}}\frac{\partial \widehat{\boldsymbol{z}}}{\partial \boldsymbol{z}} = \boldsymbol{D}(\boldsymbol{z},t)^{\mathrm{T}}. \tag{5.1.25}$$

由上面的讨论得到如下定理:

定理 5.1.2 Birkhoff 系统 (5.1.18) 在 $\mathbf{R}^{2n}$ 上的相流保持 Birkhoff 辛结构, 同时还满足约束条件 (5.1.25).

§5.2　Birkhoff 辛映射的生成函数

早在 20 世纪 80 年代冯康等人就发展了构造 Hamilton 系统辛算法的一种比较系统的方法, 即生成函数方法[Fen86,FQ87,FWQW89]. 他们用这种方法构造了标准的 Hamilton 系统, Poisson 流形上的 Hamilton 系统及一般的 $\boldsymbol{K}(\boldsymbol{z})$ 辛流形上的 Hamilton 系统的保结构算法. 对于显含时间 t 的 $\boldsymbol{K}(\boldsymbol{z},t)$ 辛流形上的 Birkhoff 系统, 在上一节已证明其相流保持 Birkhoff 辛结构 $\boldsymbol{K}(\boldsymbol{z},t)$. 本节我们将生成函数方法加以推广, 来构造一种算法——Birkhoff 辛算法或广义辛算法. 这种算法能够保持离散的 $\boldsymbol{K}(\boldsymbol{z},t)$ 结构.

5.2.1　Birkhoff 辛映射和辛格式

要利用生成函数方法构造 Birkhoff 系统的广义辛算法, 首先我们得搞清楚几个特殊映射之间的关系. 本小节先介绍 Birkhoff 辛映射 (即

$\boldsymbol{K}(\boldsymbol{z},t)$-辛映射)、梯度映射和典则映射以及 Birkhoff 辛格式的概念. 令

$$
\boldsymbol{J}_{2n} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I}_n \\ -\boldsymbol{I}_n & \boldsymbol{0} \end{pmatrix}, \qquad \boldsymbol{J}_{4n} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I}_{2n} \\ -\boldsymbol{I}_{2n} & \boldsymbol{0} \end{pmatrix},
$$
$$
\widetilde{\boldsymbol{J}}_{4n} = \begin{pmatrix} \boldsymbol{J}_{2n} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{J}_{2n} \end{pmatrix}, \quad \widetilde{\boldsymbol{K}}(\widehat{\boldsymbol{z}},\boldsymbol{z},t,t_0) = \begin{pmatrix} \boldsymbol{K}(\widehat{\boldsymbol{z}},t) & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{K}(\boldsymbol{z},t_0) \end{pmatrix}. \tag{5.2.1}
$$

定义 5.2.1 假设 U 是 $\mathbf{R}^m$ 的一个开子集. 一个包含于 $\mathbf{R}^{4n}$ 的 $2n$ 维微分流形, 记为

$$
L = \left\{ \boldsymbol{p} = \begin{pmatrix} \widehat{\boldsymbol{z}} \\ \boldsymbol{z} \end{pmatrix} \in \mathbf{R}^{4n} \,\middle|\, \boldsymbol{z} = \boldsymbol{z}(\boldsymbol{x},t_0), \widehat{\boldsymbol{z}} = \widehat{\boldsymbol{z}}(\boldsymbol{x},t), \boldsymbol{x} \in U \right\}, \tag{5.2.2}
$$

分别称为 $\boldsymbol{J}_{4n}$**-Lagrange 子流形**, $\widetilde{\boldsymbol{J}}_{4n}$**-Lagrange 子流形**, $\widetilde{\boldsymbol{K}}(\widehat{\boldsymbol{z}},\boldsymbol{z},t,t_0)$**-Lagrange 子流形**, 如果它分别具有如下不同性质:

$$
(T_{\boldsymbol{p}}L)^{\mathrm{T}} \boldsymbol{J}_{4n} (T_{\boldsymbol{p}}L) = \boldsymbol{0}, \tag{5.2.3}
$$
$$
(T_{\boldsymbol{p}}L)^{\mathrm{T}} \widetilde{\boldsymbol{J}}_{4n} (T_{\boldsymbol{p}}L) = \boldsymbol{0}, \tag{5.2.4}
$$
$$
(T_{\boldsymbol{p}}L)^{\mathrm{T}} \widetilde{\boldsymbol{K}}(\widehat{\boldsymbol{z}},\boldsymbol{z},t,t_0) (T_{\boldsymbol{p}}L) = \boldsymbol{0}, \tag{5.2.5}
$$

此处 $T_{\boldsymbol{p}}L$ 仍指微分流形 L 在点 $\boldsymbol{p}$ 处的切空间.

定义 5.2.2 记 $\boldsymbol{z} \mapsto \widehat{\boldsymbol{z}} = \boldsymbol{g}(\boldsymbol{z},t,t_0)$ 是一个从 $\mathbf{R}^{2n}$ 到它自身的, 带有参数 t 和 t_0 的映射, 它的**图**定义为

$$
\Gamma_{\boldsymbol{g}} = \left\{ \begin{pmatrix} \widehat{\boldsymbol{z}} \\ \boldsymbol{z} \end{pmatrix} \in \mathbf{R}^{4n} \,\middle|\, \widehat{\boldsymbol{z}} = \boldsymbol{g}(\boldsymbol{z},t,t_0), \widehat{\boldsymbol{z}}(t_0) = \boldsymbol{z} \in \mathbf{R}^{2n} \right\}. \tag{5.2.6}
$$

当映射 $\boldsymbol{g}$ 的图 (5.2.6) 是 $\widetilde{\boldsymbol{J}}_{4n}$-Lagrange 子流形, $\boldsymbol{J}_{4n}$-Lagrange 子流形或 $\widetilde{\boldsymbol{K}}(\widehat{\boldsymbol{z}},\boldsymbol{z},t,t_0)$- Lagrange 子流形时, 映射 $\boldsymbol{g}$ 相应地分别称为**典则映射**、**梯度映射**或 $\boldsymbol{K}(\boldsymbol{z},t)$**-辛映射**. $\boldsymbol{K}(\boldsymbol{z},t)$-辛映射也称为 **Birkhoff 辛映射**.

以上关于 Birkhoff 辛映射的定义应用起来不够直观, 为了后面能更简洁地讨论数值离散格式的辛映射性质, 我们给出另外一个 Birkhoff 辛映射的等价定义.

定义 5.2.3 设 $\boldsymbol{g}: \boldsymbol{z} \mapsto \widehat{\boldsymbol{z}} = \boldsymbol{g}(\boldsymbol{z}, t, t_0)$ 是一个带参数的可微映射, 当它满足

$$\left(\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{z}}\right)^{\mathrm{T}} \boldsymbol{K}(\boldsymbol{g}(\boldsymbol{z}, t, t_0), t) \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{z}} = \boldsymbol{K}(\boldsymbol{z}, t_0) \tag{5.2.7}$$

时, 称该映射是 **Birkhoff 辛映射**或 **$\boldsymbol{K}(\boldsymbol{z}, t)$-辛映射**.

假设 Birkhoff 系统 (5.1.1) 的相流为 $\widehat{\boldsymbol{z}} = \boldsymbol{g}^t(\boldsymbol{z}, t_0)$, 则由上一节的定理 (5.1.2) 知道 $\boldsymbol{g}^t$ 保持 $\boldsymbol{K}(\boldsymbol{z}, t)$-辛结构, 即

$$\left(\frac{\partial \boldsymbol{g}^t(\boldsymbol{z}, t_0)}{\partial \boldsymbol{z}}\right)^{\mathrm{T}} \boldsymbol{K}(\boldsymbol{g}^t(\boldsymbol{z}, t_0), t) \frac{\partial \boldsymbol{g}^t(\boldsymbol{z}, t_0)}{\partial \boldsymbol{z}} = \boldsymbol{K}(\boldsymbol{z}, t_0). \tag{5.2.8}$$

所以由定义 5.2.3 得到如下结论:

推论 5.2.1 Birkhoff 系统 (5.1.1) 的相流是 Birkhoff 辛映射.

Birkhoff 系统 (5.1.1) 相流的图可以简单表示为

$$\boldsymbol{g}^t(\boldsymbol{z}, t_0) = \boldsymbol{g}(\boldsymbol{z}, t, t_0),$$

所以它的相流组成一个 $\widetilde{\boldsymbol{K}}(\widehat{\boldsymbol{z}}, \boldsymbol{z}, t, t_0)$-Lagrange 子流形. 而标准 Hamilton 系统的相流构成一个 $\widetilde{\boldsymbol{J}}_{4n}$-Lagrange 子流形.

假定方程组 (5.1.18) 的一种离散形式可以记为

$$\boldsymbol{K}(\boldsymbol{z}^k, t_k)\partial_t^k \boldsymbol{z}^k = (\nabla_{\boldsymbol{z}} S(\boldsymbol{z}^k, t_k))_k + (\partial_t \boldsymbol{F}(\boldsymbol{z}^k, t_k))_k, \tag{5.2.9}$$

其中 $\boldsymbol{z}^k = \boldsymbol{z}(t_k)$, $t_k = t_0 + k\tau$. 这里决定离散的关键就是 ∂_t^k, 它表示 ∂_t 在第 t_k 点的离散. 假设此离散形式决定一个离散的相流

$$\boldsymbol{z}^{k+1} = \Phi(\boldsymbol{z}^k, t_k),$$

那么有如下定义:

定义 5.2.4 离散形式 (5.2.9) 称为方程组 (5.1.18) 的一个 **Birkhoff 辛格式**或 **$\boldsymbol{K}(\boldsymbol{z}, t)$-辛格式**, 如果它决定的格式 $\boldsymbol{z}^{k+1} = \Phi(\boldsymbol{z}^k, t_k)$ 保持离散的 $\boldsymbol{K}(\boldsymbol{z}, t)$-辛结构, 即

$$\left(\frac{\partial \Phi}{\partial \boldsymbol{z}^k}\right)^{\mathrm{T}} \boldsymbol{K}(\boldsymbol{z}^{k+1}, t_{k+1}) \frac{\partial \Phi}{\partial \boldsymbol{z}^k} = \boldsymbol{K}(\boldsymbol{z}^k, t_k). \tag{5.2.10}$$

5.2.2 Birkhoff 辛映射和生成函数的关系

假设在 $\mathbf{R}^{4n}$ 上存在到它自身的可逆映射, 这个映射带有参数 t 和 t_0, 比如

$$\boldsymbol{\alpha}(t,t_0):\begin{pmatrix}\widehat{\boldsymbol{z}}\\ \boldsymbol{z}\end{pmatrix}\mapsto\begin{pmatrix}\widehat{\boldsymbol{w}}\\ \boldsymbol{w}\end{pmatrix}=\begin{pmatrix}\boldsymbol{\alpha}_1(\widehat{\boldsymbol{z}},\boldsymbol{z},t,t_0)\\ \boldsymbol{\alpha}_2(\widehat{\boldsymbol{z}},\boldsymbol{z},t,t_0)\end{pmatrix},\tag{5.2.11}$$

$$\boldsymbol{\alpha}^{-1}(t,t_0):\begin{pmatrix}\widehat{\boldsymbol{w}}\\ \boldsymbol{w}\end{pmatrix}\mapsto\begin{pmatrix}\widehat{\boldsymbol{z}}\\ \boldsymbol{z}\end{pmatrix}=\begin{pmatrix}\boldsymbol{\alpha}^1(\widehat{\boldsymbol{w}},\boldsymbol{w},t,t_0)\\ \boldsymbol{\alpha}^2(\widehat{\boldsymbol{w}},\boldsymbol{w},t,t_0)\end{pmatrix}.\tag{5.2.12}$$

映射 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\alpha}^{-1}$ 的 Jacobi 矩阵分别记为

$$\begin{aligned}\boldsymbol{\alpha}_*(\widehat{\boldsymbol{z}},\boldsymbol{z},t,t_0)&=\begin{pmatrix}\boldsymbol{A}_{\boldsymbol{\alpha}} & \boldsymbol{B}_{\boldsymbol{\alpha}}\\ \boldsymbol{C}_{\boldsymbol{\alpha}} & \boldsymbol{D}_{\boldsymbol{\alpha}}\end{pmatrix},\\ \boldsymbol{\alpha}_*^{-1}(\widehat{\boldsymbol{w}},\boldsymbol{w},t,t_0)&=\begin{pmatrix}\boldsymbol{A}^{\boldsymbol{\alpha}} & \boldsymbol{B}^{\boldsymbol{\alpha}}\\ \boldsymbol{C}^{\boldsymbol{\alpha}} & \boldsymbol{D}^{\boldsymbol{\alpha}}\end{pmatrix}.\end{aligned}\tag{5.2.13}$$

假设 $\boldsymbol{\alpha}$ 是一个 $\mathbf{R}^{4n}$ 到它自身的微分同胚, 那么映射 $\boldsymbol{\alpha}$ 将一个 $\widetilde{\boldsymbol{K}}$-Lagrange 子流形映到一个 $\boldsymbol{J}_{4n}$-Lagrange 子流形, 当且仅当

$$\boldsymbol{\alpha}_*^{\mathrm{T}}\boldsymbol{J}_{4n}\boldsymbol{\alpha}_*=\widetilde{\boldsymbol{K}},\tag{5.2.14}$$

即

$$\begin{pmatrix}\boldsymbol{A}_{\boldsymbol{\alpha}} & \boldsymbol{B}_{\boldsymbol{\alpha}}\\ \boldsymbol{C}_{\boldsymbol{\alpha}} & \boldsymbol{D}_{\boldsymbol{\alpha}}\end{pmatrix}^{\mathrm{T}}\begin{pmatrix}\boldsymbol{J}_{2n} & \boldsymbol{0}\\ \boldsymbol{0} & -\boldsymbol{J}_{2n}\end{pmatrix}\begin{pmatrix}\boldsymbol{A}_{\boldsymbol{\alpha}} & \boldsymbol{B}_{\boldsymbol{\alpha}}\\ \boldsymbol{C}_{\boldsymbol{\alpha}} & \boldsymbol{D}_{\boldsymbol{\alpha}}\end{pmatrix}=\begin{pmatrix}\boldsymbol{K}(\widehat{\boldsymbol{z}},t) & \boldsymbol{0}\\ \boldsymbol{0} & -\boldsymbol{K}(\boldsymbol{z},t_0)\end{pmatrix}.\tag{5.2.15}$$

反过来, $\boldsymbol{\alpha}^{-1}$ 将一个 $\boldsymbol{J}_{4n}$-Lagrange 子流形映到一个 $\widetilde{\boldsymbol{K}}$-Lagrange 子流形.

命题 5.2.1 设 $\mathcal{M}$ 是一个包含于空间 $\mathbf{R}^{2n\times 2n}$ 的微分流形, 映射 $\boldsymbol{\alpha}$ 如上定义, 则可定义一个分数形式的变换

$$\begin{aligned}&\boldsymbol{\sigma}_{\boldsymbol{\alpha}}:\mathcal{M}\to\mathcal{M},\\ &\qquad \boldsymbol{M}\mapsto\boldsymbol{N}=\boldsymbol{\sigma}_{\boldsymbol{\alpha}}(\boldsymbol{M})=(\boldsymbol{A}_{\boldsymbol{\alpha}}\boldsymbol{M}+\boldsymbol{B}_{\boldsymbol{\alpha}})(\boldsymbol{C}_{\boldsymbol{\alpha}}\boldsymbol{M}+\boldsymbol{D}_{\boldsymbol{\alpha}})^{-1}.\end{aligned}\tag{5.2.16}$$

假设有截面条件

$$|\boldsymbol{C_\alpha M}+\boldsymbol{D_\alpha}|\neq 0 \tag{5.2.17}$$

成立, 则下面四个条件互相等价:

$$\begin{aligned}
&|\boldsymbol{C_\alpha M}+\boldsymbol{D_\alpha}|\neq 0,\\
&|\boldsymbol{MC^\alpha}-\boldsymbol{A^\alpha}|\neq 0,\\
&|\boldsymbol{C^\alpha N}+\boldsymbol{D^\alpha}|\neq 0,\\
&|\boldsymbol{NC_\alpha}-\boldsymbol{A_\alpha}|\neq 0.
\end{aligned} \tag{5.2.18}$$

这个命题的证明很简单, 这里略去. 有了上面的结果, 我们进一步得出下面的定理.

定理 5.2.1　设映射 $\boldsymbol{\alpha}$ 如上定义. 假定 $\boldsymbol{z}\mapsto\widehat{\boldsymbol{z}}=\boldsymbol{g}(\boldsymbol{z},t,t_0)$ 是定义在 $\mathbf{R}^{2n}$ 上的某一区域 $\widetilde{\mathcal{R}}$ 上的一个 $\boldsymbol{K}(\boldsymbol{z},t)$-辛映射, 它的 Jacobi 矩阵 $\boldsymbol{g_z}(\boldsymbol{z},t,t_0)$ 记为 $\boldsymbol{M}(\boldsymbol{z},t,t_0)$. 如果 Jacobi 矩阵 $\boldsymbol{M}(\boldsymbol{z},t,t_0)$ 在区域 $\widetilde{\mathcal{R}}$ 上满足截面条件

$$|\boldsymbol{C_\alpha}(\boldsymbol{g}(\boldsymbol{z},t,t_0),\boldsymbol{z},t,t_0)\boldsymbol{M}(\boldsymbol{z},t,t_0)+\boldsymbol{D_\alpha}(\boldsymbol{g}(\boldsymbol{z},t,t_0),\boldsymbol{z},t,t_0)|\neq 0, \tag{5.2.19}$$

则在区域 $\widetilde{\mathcal{R}}$ 上存在唯一的梯度映射 $\boldsymbol{w}\mapsto\widehat{\boldsymbol{w}}=\boldsymbol{f}(\boldsymbol{w},t,t_0)$, 以及存在标量函数 (即生成函数) $\phi(\boldsymbol{w},t,t_0)$, 使得

$$\boldsymbol{f}(\boldsymbol{w},t,t_0)=\phi_{\boldsymbol{w}}(\boldsymbol{w},t,t_0), \tag{5.2.20}$$

$$\begin{aligned}
\boldsymbol{\alpha}_1(\boldsymbol{g}(\boldsymbol{z},t,t_0),\boldsymbol{z},t,t_0)&=\boldsymbol{f}(\boldsymbol{\alpha}_2(\boldsymbol{g}(\boldsymbol{z},t,t_0),\boldsymbol{z},t,t_0),t,t_0)\\
&=\phi_{\boldsymbol{w}}(\boldsymbol{\alpha}_2(\boldsymbol{g}(\boldsymbol{z},t,t_0),\boldsymbol{z},t,t_0),t,t_0),
\end{aligned} \tag{5.2.21}$$

其中 $\boldsymbol{\alpha}_1$ 关于变量 $\boldsymbol{z}$ 和 t 唯一确定, 并且 $\boldsymbol{f}(\boldsymbol{w},t,t_0)$ 的 Jacobi 矩阵 $\boldsymbol{N}(\boldsymbol{w},t,t_0)=\boldsymbol{f_w}(\boldsymbol{w},t,t_0)$ 满足

$$\boldsymbol{N}=(\boldsymbol{A_\alpha M}+\boldsymbol{B_\alpha})(\boldsymbol{C_\alpha M}+\boldsymbol{D_\alpha})^{-1}, \tag{5.2.22}$$

$$\boldsymbol{M}=(\boldsymbol{A^\alpha N}+\boldsymbol{B^\alpha})(\boldsymbol{C^\alpha N}+\boldsymbol{D^\alpha})^{-1}. \tag{5.2.23}$$

证明 映射 $\boldsymbol{g}$ 的图 $\Gamma_{\boldsymbol{g}}$ 通过变换 $\boldsymbol{\alpha}$ 映到下面的集合:

$$\boldsymbol{\alpha}(\Gamma_{\boldsymbol{g}}) = \left\{ \begin{pmatrix} \widehat{\boldsymbol{w}} \\ \boldsymbol{w} \end{pmatrix} \in \mathbf{R}^{4n} \middle| \begin{array}{l} \widehat{\boldsymbol{w}} = \boldsymbol{\alpha}_1(\boldsymbol{g}(\boldsymbol{z}, t, t_0), \boldsymbol{z}, t, t_0), \\ \boldsymbol{w} = \boldsymbol{\alpha}_2(\boldsymbol{g}(\boldsymbol{z}, t, t_0), \boldsymbol{z}, t, t_0) \end{array} \right\}. \quad (5.2.24)$$

根据截面条件 (5.2.19), 得到

$$\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{z}} = \frac{\partial \boldsymbol{\alpha}_2}{\partial \widehat{\boldsymbol{z}}} \cdot \frac{\partial \widehat{\boldsymbol{z}}}{\partial \boldsymbol{z}} + \frac{\partial \boldsymbol{\alpha}_2}{\partial \boldsymbol{z}} = \boldsymbol{C}_{\boldsymbol{\alpha}} \boldsymbol{M} + \boldsymbol{D}_{\boldsymbol{\alpha}} \neq \boldsymbol{0}, \quad (5.2.25)$$

所以函数 $\boldsymbol{w} = \boldsymbol{\alpha}_2(\boldsymbol{g}(\boldsymbol{z}, t, t_0), \boldsymbol{z}, t, t_0)$ 是可逆的, 它的逆函数记为 $\boldsymbol{z} = \boldsymbol{z}(\boldsymbol{w}, t, t_0)$. 令

$$\widehat{\boldsymbol{w}} = \boldsymbol{f}(\boldsymbol{w}, t, t_0) = \boldsymbol{\alpha}_1(\boldsymbol{g}(\boldsymbol{z}, t, t_0), \boldsymbol{z}, t, t_0)\big|_{\boldsymbol{z}=\boldsymbol{z}(\boldsymbol{w},t,t_0)}, \quad (5.2.26)$$

那么

$$\begin{aligned} \boldsymbol{N} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{w}} &= \left(\frac{\partial \boldsymbol{\alpha}_1}{\partial \widehat{\boldsymbol{z}}} \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{z}} + \frac{\partial \boldsymbol{\alpha}_1}{\partial \boldsymbol{z}} \right) \frac{\partial \boldsymbol{z}}{\partial \boldsymbol{w}} \\ &= (\boldsymbol{A}_{\boldsymbol{\alpha}} \boldsymbol{M} + \boldsymbol{B}_{\boldsymbol{\alpha}})(\boldsymbol{C}_{\boldsymbol{\alpha}} \boldsymbol{M} + \boldsymbol{D}_{\boldsymbol{\alpha}})^{-1}, \end{aligned} \quad (5.2.27)$$

即 (5.2.22) 式成立, 从而 (5.2.23) 式也成立. 注意到 $\boldsymbol{\alpha}(\Gamma_{\boldsymbol{g}})$ 在点 $\boldsymbol{z}$ 处的切空间是

$$T_{\boldsymbol{z}}(\boldsymbol{\alpha}(\Gamma_{\boldsymbol{g}})) = \begin{pmatrix} \dfrac{\partial \widehat{\boldsymbol{w}}}{\partial \boldsymbol{z}} \\ \dfrac{\partial \boldsymbol{w}}{\partial \boldsymbol{z}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}_{\boldsymbol{\alpha}} \boldsymbol{M} + \boldsymbol{B}_{\boldsymbol{\alpha}} \\ \boldsymbol{C}_{\boldsymbol{\alpha}} \boldsymbol{M} + \boldsymbol{D}_{\boldsymbol{\alpha}} \end{pmatrix}. \quad (5.2.28)$$

因为 $\boldsymbol{z} \mapsto \widehat{\boldsymbol{z}}$ 是 $\boldsymbol{K}(\boldsymbol{z}, t)$-辛映射, 所以

$$\begin{aligned} \boldsymbol{0} &= (\boldsymbol{M}^{\mathrm{T}}, \boldsymbol{I}) \widetilde{\boldsymbol{K}} \begin{pmatrix} \boldsymbol{M} \\ \boldsymbol{I} \end{pmatrix} \\ &= (\boldsymbol{M}^{\mathrm{T}}, \boldsymbol{I}) \boldsymbol{\alpha}_*^{\mathrm{T}} \boldsymbol{J}_{4n} \boldsymbol{\alpha}_* \begin{pmatrix} \boldsymbol{M} \\ \boldsymbol{I} \end{pmatrix} \\ &= ((\boldsymbol{A}_{\boldsymbol{\alpha}} \boldsymbol{M} + \boldsymbol{B}_{\boldsymbol{\alpha}})^{\mathrm{T}}, (\boldsymbol{C}_{\boldsymbol{\alpha}} \boldsymbol{M} + \boldsymbol{D}_{\boldsymbol{\alpha}})^{\mathrm{T}}) \boldsymbol{J}_{4n} \begin{pmatrix} \boldsymbol{A}_{\boldsymbol{\alpha}} \boldsymbol{M} + \boldsymbol{B}_{\boldsymbol{\alpha}} \\ \boldsymbol{C}_{\boldsymbol{\alpha}} \boldsymbol{M} + \boldsymbol{D}_{\boldsymbol{\alpha}} \end{pmatrix} \\ &= T_{\boldsymbol{z}}(\boldsymbol{\alpha}(\Gamma_{\boldsymbol{g}}))^{\mathrm{T}} \boldsymbol{J}_{4n} T_{\boldsymbol{z}}(\boldsymbol{\alpha}(\Gamma_{\boldsymbol{g}})), \end{aligned} \quad (5.2.29)$$

所以 $\boldsymbol{\alpha}(\Gamma_{\boldsymbol{g}})$ 是一个 $\boldsymbol{J}_{4n}$-Lagrange 子流形. 而且, 由上面的运算知

$$(\boldsymbol{A_\alpha M}+\boldsymbol{B_\alpha})^{\mathrm{T}}(\boldsymbol{C_\alpha M}+\boldsymbol{D_\alpha})-(\boldsymbol{C_\alpha M}+\boldsymbol{D_\alpha})^{\mathrm{T}}(\boldsymbol{A_\alpha M}+\boldsymbol{B_\alpha})=\boldsymbol{0}, \tag{5.2.30}$$

即 Jacobi 矩阵 $\boldsymbol{N}=(\boldsymbol{A_\alpha M}+\boldsymbol{B_\alpha})(\boldsymbol{C_\alpha M}+\boldsymbol{D_\alpha})^{-1}$ 是对称的. 这意味着映射 $\widehat{\boldsymbol{w}}=\boldsymbol{f}(\boldsymbol{w},t,t_0)$ 是一个梯度映射. 根据著名的 Poincaré 引理, 存在一个标量函数 $\phi(\boldsymbol{w},t,t_0)$, 使得

$$\boldsymbol{f}(\boldsymbol{w},t,t_0)=\phi_{\boldsymbol{w}}(\boldsymbol{w},t,t_0). \tag{5.2.31}$$

又 $\boldsymbol{z}(\boldsymbol{w},t,t_0)\circ\boldsymbol{\alpha}_2(\boldsymbol{g}(\boldsymbol{z},t,t_0),\boldsymbol{z},t,t_0)\equiv\boldsymbol{z}$, 将 $\boldsymbol{w}=\boldsymbol{\alpha}_2(\boldsymbol{g}(\boldsymbol{z},t,t_0),\boldsymbol{z},t,t_0)$ 代入方程 (5.2.26) 和 (5.2.31), 通过简单的计算就可以得到方程 (5.2.21). □

推论 5.2.2　如果 $\boldsymbol{f}(\boldsymbol{w},t,t_0)$ 是定理 5.2.1 中得到的函数, 那么它也是下面的方程的解:

$$\boldsymbol{\alpha}^1(\boldsymbol{f}(\boldsymbol{w},t,t_0),\boldsymbol{w},t,t_0)=\boldsymbol{g}(\boldsymbol{\alpha}^2(\boldsymbol{f}(\boldsymbol{w},t,t_0),\boldsymbol{w},t,t_0),t,t_0). \tag{5.2.32}$$

推论的证明很直接, 略去.

定理 5.2.2　设 $\boldsymbol{\alpha}$ 是命题 5.2.1 中定义的映射, 又设 $\boldsymbol{w}\mapsto\widehat{\boldsymbol{w}}=\boldsymbol{f}(\boldsymbol{w},t,t_0)$ 是定义在空间 $\mathbf{R}^{2n}$ 的某一区域 $\widetilde{\mathcal{R}}$ 上的梯度映射, 它有 Jacobi 矩阵 $\boldsymbol{f}_{\boldsymbol{w}}(\boldsymbol{w},t,t_0)=\boldsymbol{N}(\boldsymbol{w},t,t_0)$. 如果在区域 $\widetilde{\mathcal{R}}$ 上, $\boldsymbol{N}$ 满足截面条件

$$|\boldsymbol{C}^{\boldsymbol{\alpha}}(\boldsymbol{f}(\boldsymbol{w},t,t_0),\boldsymbol{w},t,t_0)\boldsymbol{N}(\boldsymbol{w},t,t_0)+\boldsymbol{D}^{\boldsymbol{\alpha}}(\boldsymbol{f}(\boldsymbol{w},t,t_0),\boldsymbol{w},t,t_0)|\neq 0, \tag{5.2.33}$$

那么在区域 $\widetilde{\mathcal{R}}$ 上存在唯一的 $\boldsymbol{K}(\boldsymbol{z},t)$-辛映射 $\boldsymbol{z}\mapsto\widehat{\boldsymbol{z}}=\boldsymbol{g}(\boldsymbol{z},t,t_0)$, 它有 Jacobi 矩阵 $\boldsymbol{g}_{\boldsymbol{z}}(\boldsymbol{z},t,t_0)=\boldsymbol{M}(\boldsymbol{z},t,t_0)$, 使得

$$\begin{aligned}&\boldsymbol{\alpha}^1(\boldsymbol{f}(\boldsymbol{w},t,t_0),\boldsymbol{w},t,t_0)=\boldsymbol{g}(\boldsymbol{\alpha}^2(\boldsymbol{f}(\boldsymbol{w},t,t_0),\boldsymbol{w},t,t_0),t,t_0),\\&\boldsymbol{M}=(\boldsymbol{A}^{\boldsymbol{\alpha}}\boldsymbol{N}+\boldsymbol{B}^{\boldsymbol{\alpha}})(\boldsymbol{C}^{\boldsymbol{\alpha}}\boldsymbol{N}+\boldsymbol{D}^{\boldsymbol{\alpha}})^{-1},\\&\boldsymbol{N}=(\boldsymbol{A_\alpha M}+\boldsymbol{B_\alpha})(\boldsymbol{C_\alpha M}+\boldsymbol{D_\alpha})^{-1},\end{aligned} \tag{5.2.34}$$

并且 $\boldsymbol{g}(\boldsymbol{z},t,t_0)$ 是隐式方程

$$\boldsymbol{\alpha}_1(\boldsymbol{g}(\boldsymbol{z},t,t_0),\boldsymbol{z},t,t_0)=\boldsymbol{f}(\boldsymbol{\alpha}_2(\boldsymbol{g}(\boldsymbol{z},t,t_0),\boldsymbol{z},t,t_0),\boldsymbol{z},t,t_0) \tag{5.2.35}$$

的解.

这个定理的证明和定理 5.2.1 的证明类似, 略去.

§5.3 Birkhoff 方程的 $\boldsymbol{K}(\boldsymbol{z},t)$-辛差分格式

在前两节中, 我们就 Birkhoff 系统特有的 $\boldsymbol{K}(\boldsymbol{z},t)$-辛结构, 讨论了它的相流所表现出的保结构特性以及它的相流和生成函数之间的关系. 在这一节中, 我们将通过逼近生成函数给出具体的构造 Birkhoff 辛差分格式的步骤.

5.3.1 Birkhoff 系统的相流的生成函数

我们用 $\boldsymbol{g}^t(\boldsymbol{z},t_0)$ 记 Birkhoff 系统的相流, 后面的讨论都是在局部空间 $\{(\boldsymbol{z},t)\}$ 上进行的, 所以相流 $\boldsymbol{g}^t(\boldsymbol{z},t_0)$ 是 $\boldsymbol{K}(\boldsymbol{z},t)$-辛映射的一个单参数群, 即

$$\boldsymbol{g}^{t_0}=\mathrm{id},\quad \boldsymbol{g}^{t_1+t_2}=\boldsymbol{g}^{t_1}\cdot\boldsymbol{g}^{t_2}, \tag{5.3.1}$$

这里认为 $\boldsymbol{z}$ 是 $\widehat{\boldsymbol{z}}$ 在 $t=t_0$ 时的初值, 而 $\widehat{\boldsymbol{z}}(\boldsymbol{z},t,t_0)=\boldsymbol{g}^t(\boldsymbol{z},t_0)\triangleq \boldsymbol{g}(t;\boldsymbol{z},t_0)$ 是 Birkhoff 系统 (5.1.1) 在这个初始条件下的解. 以下定理就给出了 Birkhoff 相流的生成函数.

定理 5.3.1 设映射 $\boldsymbol{\alpha}$ 如命题 5.2.1 所定义, 映射 $\boldsymbol{z}\mapsto\widehat{\boldsymbol{z}}=\boldsymbol{g}(t;\boldsymbol{z},t_0)$ 是 Birkhoff 系统 (5.1.1) 的相流, $\boldsymbol{M}(t;\boldsymbol{z},t_0)=\boldsymbol{g}_{\boldsymbol{z}}(t;\boldsymbol{z},t_0)$ 是它的 Jacobi 矩阵. 如果在初值 $t=t_0,\widehat{\boldsymbol{z}}=\boldsymbol{z}$ 处, 有

$$|\boldsymbol{C}_{\boldsymbol{\alpha}}(\boldsymbol{z},\boldsymbol{z},t_0,t_0)+\boldsymbol{D}_{\boldsymbol{\alpha}}(\boldsymbol{z},\boldsymbol{z},t_0,t_0)|\neq 0, \tag{5.3.2}$$

那么当 $|t-t_0|$ 足够小时, 存在初值 $\boldsymbol{z}$ 的邻域 $\mathcal{R}$, 在这个邻域内, 存在梯度映射 $\boldsymbol{w}\mapsto\widehat{\boldsymbol{w}}=\boldsymbol{f}(\boldsymbol{w},t,t_0)$, 它的 Jacobi 矩阵 $\boldsymbol{f}_{\boldsymbol{w}}(\boldsymbol{w},t,t_0)=\boldsymbol{N}(\boldsymbol{w},t,t_0)$ 是对称的, 还存在一个标量函数, 即生成函数 $\phi(\boldsymbol{w},t,t_0)$, 使得

$$\boldsymbol{f}(\boldsymbol{w},t,t_0)=\phi_{\boldsymbol{w}}(\boldsymbol{w},t,t_0), \tag{5.3.3}$$

$$\frac{\partial}{\partial t}\phi_{\boldsymbol{w}}(\boldsymbol{w},t,t_0)=\mathcal{A}(\phi_{\boldsymbol{w}}(\boldsymbol{w},t,t_0),\boldsymbol{w},\phi_{\boldsymbol{ww}}(\boldsymbol{w},t,t_0),t,t_0), \tag{5.3.4}$$

$$\mathcal{A}\left(\widehat{\boldsymbol{w}},\boldsymbol{w},\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}},t,t_0\right)=\bar{\mathcal{A}}\left(\widehat{\boldsymbol{z}}(\widehat{\boldsymbol{w}},\boldsymbol{w},t,t_0),\boldsymbol{z}(\widehat{\boldsymbol{w}},\boldsymbol{w},t,t_0),\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}},t,t_0\right),\tag{5.3.5}$$

$$\begin{aligned}\bar{\mathcal{A}}\left(\widehat{\boldsymbol{z}},\boldsymbol{z},\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}},t,t_0\right)&=\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\boldsymbol{w}}(\widehat{\boldsymbol{z}},\boldsymbol{z},t,t_0)-\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{w}(\widehat{\boldsymbol{z}},\boldsymbol{z},t,t_0)\\&=\left(\boldsymbol{A}_{\boldsymbol{\alpha}}-\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\boldsymbol{C}_{\boldsymbol{\alpha}}\right)\boldsymbol{K}^{-1}\boldsymbol{D}(\widehat{\boldsymbol{z}},t)\\&\quad+\frac{\partial\boldsymbol{\alpha}_1}{\partial t}-\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\frac{\partial\boldsymbol{\alpha}_2}{\partial t},\end{aligned}\tag{5.3.6}$$

$$\begin{aligned}\boldsymbol{\alpha}_1(\boldsymbol{g}(t;\boldsymbol{z},t_0),\boldsymbol{z},t,t_0)&=\boldsymbol{f}(\boldsymbol{\alpha}_2(\boldsymbol{g}(t;\boldsymbol{z},t_0),\boldsymbol{z},t,t_0),t,t_0)\\&=\phi_{\boldsymbol{w}}(\boldsymbol{\alpha}_2(\boldsymbol{g}(t;\boldsymbol{z},t_0),\boldsymbol{z},t,t_0),t,t_0),\end{aligned}\tag{5.3.7}$$

$\boldsymbol{\alpha}_1$ 关于 $(\boldsymbol{z},t_0)$ 唯一确定, 且

$$\begin{aligned}\boldsymbol{N}&=\boldsymbol{\sigma}_{\boldsymbol{\alpha}}(\boldsymbol{M})=(\boldsymbol{A}_{\boldsymbol{\alpha}}\boldsymbol{M}+\boldsymbol{B}_{\boldsymbol{\alpha}})(\boldsymbol{C}_{\boldsymbol{\alpha}}\boldsymbol{M}+\boldsymbol{D}_{\boldsymbol{\alpha}})^{-1},\\\boldsymbol{M}&=\boldsymbol{\sigma}_{\boldsymbol{\alpha}^{-1}}(\boldsymbol{N})=(\boldsymbol{A}^{\boldsymbol{\alpha}}\boldsymbol{N}+\boldsymbol{B}^{\boldsymbol{\alpha}})(\boldsymbol{C}^{\boldsymbol{\alpha}}\boldsymbol{N}+\boldsymbol{D}^{\boldsymbol{\alpha}})^{-1}.\end{aligned}\tag{5.3.8}$$

证明　假设 $\boldsymbol{M}(t;\boldsymbol{z},t_0)$ 关于 $\boldsymbol{z}$ 和 t 可微. 由条件 (5.3.2) 知道, 对于足够小的 $|t-t_0|$, 在初值 $\boldsymbol{z}\in\mathbf{R}^{2n}$ 的某个邻域, 有

$$\left|\boldsymbol{C}_{\boldsymbol{\alpha}}(\widehat{\boldsymbol{z}},\boldsymbol{z},t,t_0)\boldsymbol{M}(t;\boldsymbol{z},t_0)+\boldsymbol{D}_{\boldsymbol{\alpha}}(\widehat{\boldsymbol{z}},\boldsymbol{z},t,t_0)\right|\neq 0.\tag{5.3.9}$$

另外, 由于 Birkhoff 系统相流是一个辛映射, 因此由定理 5.2.1 知, 存在一个含有 t 参数的梯度映射 $\widehat{\boldsymbol{w}}=\boldsymbol{f}(\boldsymbol{w},t,t_0)$ 和一个标量函数 $\phi(\boldsymbol{w},t,t_0)$, 使得

$$\boldsymbol{f}(\boldsymbol{w},t,t_0)=\phi_{\boldsymbol{w}}(\boldsymbol{w},t,t_0),\tag{5.3.10}$$

$$\frac{\partial}{\partial t}\boldsymbol{f}(\boldsymbol{w},t,t_0)=\frac{\partial}{\partial t}\phi_{\boldsymbol{w}}(\boldsymbol{w},t,t_0).\tag{5.3.11}$$

注意到 $\widehat{\boldsymbol{z}}=\boldsymbol{g}(t;\boldsymbol{z},t_0)$ 是初值问题

$$\begin{cases}\dfrac{\mathrm{d}\widehat{\boldsymbol{z}}}{\mathrm{d}t}=\boldsymbol{K}^{-1}(\widehat{\boldsymbol{z}},t)\left(\nabla B+\dfrac{\partial\boldsymbol{F}}{\partial t}\right)(\widehat{\boldsymbol{z}},t),\\ \widehat{\boldsymbol{z}}\big|_{t=t_0}=\boldsymbol{z}\end{cases}\tag{5.3.12}$$

的解, 因此从等式

$$\widehat{\boldsymbol{w}}=\boldsymbol{\alpha}_1(\widehat{\boldsymbol{z}},\boldsymbol{z},t,t_0),\quad \boldsymbol{w}=\boldsymbol{\alpha}_2(\widehat{\boldsymbol{z}},\boldsymbol{z},t,t_0)\tag{5.3.13}$$

可以得到

$$\begin{aligned}\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t} &= \frac{\partial\widehat{\boldsymbol{w}}}{\partial\widehat{\boldsymbol{z}}}\frac{\mathrm{d}\widehat{\boldsymbol{z}}}{\mathrm{d}t} + \frac{\partial}{\partial t}\boldsymbol{\alpha}_1(\widehat{\boldsymbol{z}},\boldsymbol{z},t,t_0)\\ &= \boldsymbol{A}_{\boldsymbol{\alpha}}\boldsymbol{K}^{-1}\left(\nabla B + \frac{\partial \boldsymbol{F}}{\partial t}\right) + \frac{\partial\boldsymbol{\alpha}_1}{\partial t},\\ \frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t} &= \boldsymbol{C}_{\boldsymbol{\alpha}}\boldsymbol{K}^{-1}\left(\nabla B + \frac{\partial \boldsymbol{F}}{\partial t}\right) + \frac{\partial\boldsymbol{\alpha}_2}{\partial t}.\end{aligned} \tag{5.3.14}$$

所以有

$$\begin{aligned}\frac{\partial\widehat{\boldsymbol{w}}}{\partial t} &= \frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t} - \frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\\ &= \left(\boldsymbol{A}_{\boldsymbol{\alpha}} - \frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\boldsymbol{C}_{\boldsymbol{\alpha}}\right)\boldsymbol{K}^{-1}\left(\nabla B + \frac{\partial \boldsymbol{F}}{\partial t}\right) + \frac{\partial\boldsymbol{\alpha}_1}{\partial t} - \frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\frac{\partial\boldsymbol{\alpha}_2}{\partial t}.\end{aligned} \tag{5.3.15}$$

又因为 $\dfrac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}} \neq \boldsymbol{0}$, 所以 $\boldsymbol{w} = \boldsymbol{w}(\widehat{\boldsymbol{w}},t)$ 是存在且可解的, 但需要注意的是它并不能直接由变换 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\alpha}^{-1}$ 解出. 令

$$\bar{\mathcal{A}}\left(\widehat{\boldsymbol{z}},\boldsymbol{z},\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}},t,t_0\right) = \frac{\partial\widehat{\boldsymbol{w}}}{\partial t}, \tag{5.3.16}$$

记

$$\mathcal{A}\left(\widehat{\boldsymbol{w}},\boldsymbol{w},\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}},t,t_0\right) = \bar{\mathcal{A}}\left(\widehat{\boldsymbol{z}}(\widehat{\boldsymbol{w}},\boldsymbol{w},t,t_0),\boldsymbol{z}(\widehat{\boldsymbol{w}},\boldsymbol{w},t,t_0),\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}},t,t_0\right), \tag{5.3.17}$$

那么从方程 (5.3.11) 可以得到

$$\frac{\partial}{\partial t}\phi_{\boldsymbol{w}}(\boldsymbol{w},t,t_0) = \mathcal{A}(\phi_{\boldsymbol{w}},\boldsymbol{w},\phi_{\boldsymbol{ww}},t,t_0). \tag{5.3.18}$$

□

以上定理和讨论都是针对非自治的 Birkhoff 系统给出的, 对于自治和半自治的 Birkhoff 系统也有基本相同的结论, 所以以上结论也可以包括自治和半自治的情形. 我们前面已经说明了自治和半自治的 Birkhoff 系统就是通常所说一般辛结构下的 Hamilton 系统. 因此 Birkhoff 系统也是 Hamilton 系统的一种推广. 但是需要说明的是, 非自治 Birkhoff 系统除了辛结构显含时间变量 t 之外, 它还有本质上区

别于 Hamilton 系统的一点, 即有外力项. 自治和半自治的 Birkhoff 系统 (即一般辛结构下的 Hamilton 系统) 因为没有外力项, 所以其生成函数方法比较简明. 冯康和汪道柳在文献 [Wan91] 中详细地讨论了一般辛结构和 Poisson 结构下 Hamilton 系统的生成函数方法. 在这里, 简单说明一下自治和半自治的 Birkhoff 系统的生成函数方法与非自治的 Birkhoff 系统的生成函数方法的不同之处: 前者因为辛结构显含时间变量 t, 所以映射 $\boldsymbol{\alpha}$ 不含参数 t, 因此有

$$
\begin{aligned}
\frac{\partial \widehat{\boldsymbol{w}}}{\partial t} &= \frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t} - \frac{\partial \widehat{\boldsymbol{w}}}{\partial \boldsymbol{w}}\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t} = \left(\boldsymbol{A}_\alpha - \frac{\partial \widehat{\boldsymbol{w}}}{\partial \boldsymbol{w}}\boldsymbol{C}_{\boldsymbol{\alpha}}\right)\boldsymbol{K}^{-1}\nabla B \\
&= -\left(\boldsymbol{B}^{\boldsymbol{\alpha}\mathrm{T}} + \left(\frac{\partial \widehat{\boldsymbol{w}}}{\partial \boldsymbol{w}}\right)^{\mathrm{T}}\boldsymbol{A}^{\boldsymbol{\alpha}\mathrm{T}}\right)\nabla_{\boldsymbol{z}} B \\
&= -B_{\boldsymbol{w}}(\widehat{\boldsymbol{z}}(\widehat{\boldsymbol{w}},\boldsymbol{w})) \quad (\text{或} = -B_{\boldsymbol{w}}(\widehat{\boldsymbol{z}}(\widehat{\boldsymbol{w}},\boldsymbol{w}),t)),
\end{aligned}
\tag{5.3.19}
$$

进而得到自治情形下的 Hamilton-Jacobi 方程

$$
\frac{\partial \phi(\boldsymbol{w},t)}{\partial t} = -B(\widehat{\boldsymbol{z}}(\phi_{\boldsymbol{w}},\boldsymbol{w})),
\tag{5.3.20}
$$

或者是半自治情形的 Hamilton-Jacobi 方程

$$
\frac{\partial \phi(\boldsymbol{w},t,t_0)}{\partial t} = -B(\widehat{\boldsymbol{z}}(\phi_{\boldsymbol{w}},\boldsymbol{w}),t).
\tag{5.3.21}
$$

所以, 我们在给出非自治的 Birkhoff 系统的生成函数方法时, 因为外力项而无法直接得到关于生成函数 $\phi(\boldsymbol{w},t,t_0)$ 的 Hamilton-Jacobi 方程, 但是我们得到了关于 $\phi_{\boldsymbol{w}}(\boldsymbol{w},t,t_0)$ 的 Hamilton-Jacobi 方程 (5.3.18). 而构造 Birkhoff 系统的辛格式恰好只需要 $\phi_{\boldsymbol{w}}(\boldsymbol{w},t,t_0)$ 的各阶关于 t 的偏导数, 不需要 $\phi(\boldsymbol{w},t,t_0)$, 可以说在我们的理论中是以 $\phi_{\boldsymbol{w}}(\boldsymbol{w},t,t_0)$ 作为生成函数的.

5.3.2 构造 Birkhoff 辛差分格式

下面首先给出定理 5.3.1 中出现的函数 $\mathcal{A}(\phi_{\boldsymbol{w}},\boldsymbol{w},\phi_{\boldsymbol{ww}},t,t_0)$ 关于时间 t 的全导数. 在具体的算例中, 我们总会求出函数 $\mathcal{A}$, 要求出它的全导数就很容易, 所以为了后面定理的证明, 这里给出的只是一个比较

抽象的记号, 很容易看懂, 也不影响结论的正确性:

$$\begin{aligned}
\mathrm{D}_t^k\mathcal{A} &= \partial_{\phi_{\boldsymbol{w}}}\mathcal{A}\left(\sum_{i=0}^{\infty}(t-t_0)^i\phi_{\boldsymbol{w}}^{(k+i)}\right)+\partial_{\phi_{\boldsymbol{ww}}}\mathcal{A}\left(\sum_{i=0}^{\infty}(t-t_0)^i\phi_{\boldsymbol{ww}}^{(k+i)}\right)\\
&+\partial_t\partial_{\phi_{\boldsymbol{w}}}\mathcal{A}\left(\sum_{i=0}^{\infty}(t-t_0)^i\phi_{\boldsymbol{w}}^{(k-1+i)}\right)\\
&+\partial_t\partial_{\phi_{\boldsymbol{ww}}}\mathcal{A}\left(\sum_{i=0}^{\infty}(t-t_0)^i\phi_{\boldsymbol{ww}}^{(k-1+i)}\right)\\
&+\sum_{m=0}^{k}\mathrm{C}_k^m\sum_{n=1}^{k-m}\mathrm{C}_{k-m}^n\sum_{l=1}^{k-m-n}\sum_{\substack{h_1+\cdots+h_n+j_1\\+\cdots+j_l=k-m}}\partial_{\phi_{\boldsymbol{w}}}^n\partial_{\phi_{\boldsymbol{ww}}}^l\\
&\cdot\partial_t^m\mathcal{A}\Bigg(\sum_{i=0}^{\infty}(t-t_0)^i\phi_{\boldsymbol{w}}^{(h_1+i)},\cdots,\sum_{i=0}^{\infty}(t-t_0)^i\phi_{\boldsymbol{w}}^{(h_n+i)},\\
&\qquad\sum_{i=0}^{\infty}(t-t_0)^i\phi_{\boldsymbol{ww}}^{(j_1+i)},\cdots,\sum_{i=0}^{\infty}(t-t_0)^i\phi_{\boldsymbol{w}}^{(j_l+i)}\Bigg).
\end{aligned}\tag{5.3.22}$$

在初始点 $t=t_0$ 处, 函数 $\mathcal{A}$ 的全导数就是

$$\begin{aligned}
&\mathrm{D}_t^k\mathcal{A}(\phi_{\boldsymbol{w}}^{(0)},\boldsymbol{w},\phi_{\boldsymbol{ww}}^{(0)},t_0,t_0)\\
&=\partial_{\phi_{\boldsymbol{w}}}\mathcal{A}\phi_{\boldsymbol{w}}^{(k)}+\partial_{\phi_{\boldsymbol{ww}}}\mathcal{A}\phi_{\boldsymbol{ww}}^{(k)}+\partial_t\partial_{\phi_{\boldsymbol{w}}}\mathcal{A}\phi_{\boldsymbol{w}}^{(k-1)}+\partial_t\partial_{\phi_{\boldsymbol{ww}}}\mathcal{A}\phi_{\boldsymbol{ww}}^{(k-1)}\\
&\quad+\sum_{m=0}^{k}\mathrm{C}_k^m\sum_{n=1}^{k-m}\mathrm{C}_{k-m}^n\sum_{l=1}^{k-m-n}\sum_{\substack{h_1+\cdots+h_n+j_1\\+\cdots+j_l=k-m}}\partial_{\phi_{\boldsymbol{w}}}^n\partial_{\phi_{\boldsymbol{ww}}}^l\\
&\quad\cdot\partial_t^m\mathcal{A}(\phi_{\boldsymbol{w}}^{(0)},\boldsymbol{w},\phi_{\boldsymbol{ww}}^{(0)},t_0,t_0)(\phi_{\boldsymbol{w}}^{(h_1)},\cdots,\phi_{\boldsymbol{w}}^{(h_n)},\phi_{\boldsymbol{ww}}^{(j_1)},\cdots,\phi_{\boldsymbol{ww}}^{(j_l)}).
\end{aligned}\tag{5.3.23}$$

定理 5.3.2 假设 $\mathcal{A}$ 和 $\boldsymbol{\alpha}$ 是解析的, 那么当 $|t-t_0|$ 足够小时, 生成函数 $\phi_{\boldsymbol{w}}(\boldsymbol{w},t,t_0)$ 可以展成级数:

$$\phi_{\boldsymbol{w}}(\boldsymbol{w},t,t_0)=\sum_{k=0}^{\infty}(t-t_0)^k\phi_{\boldsymbol{w}}^{(k)}(\boldsymbol{w},t_0),\tag{5.3.24}$$

并且是收敛的. 对于任意的 $k \geqslant 0$, $\phi_{\boldsymbol{w}}^{(k)}$ 可以由下面的式子递推决定:

$$\phi_{\boldsymbol{w}}^{(0)}(\boldsymbol{w}, t_0) = \boldsymbol{f}(\boldsymbol{w}, t_0, t_0), \tag{5.3.25}$$

$$\phi_{\boldsymbol{w}}^{(1)}(\boldsymbol{w}, t_0) = \mathcal{A}(\phi_{\boldsymbol{w}}^{(0)}, \boldsymbol{w}, \phi_{\boldsymbol{ww}}^{(0)}, t_0, t_0), \tag{5.3.26}$$

$$\phi_{\boldsymbol{w}}^{(k+1)}(\boldsymbol{w}, t_0) = \frac{1}{(k+1)!}\mathrm{D}_t^k \mathcal{A}(\phi_{\boldsymbol{w}}^{(0)}, \boldsymbol{w}, \phi_{\boldsymbol{ww}}^{(0)}, t_0, t_0). \tag{5.3.27}$$

证明　对 (5.3.24) 式关于 $\boldsymbol{w}$ 和 t 作微分, 得到

$$\phi_{\boldsymbol{ww}}(\boldsymbol{w}, t, t_0) = \sum_{k=0}^{\infty}(t - t_0)^k \phi_{\boldsymbol{ww}}^{(k)}(\boldsymbol{w}, t_0) \tag{5.3.28}$$

和

$$\frac{\partial}{\partial t}\phi_{\boldsymbol{w}}(\boldsymbol{w}, t, t_0) = \sum_{k=0}^{\infty}(k+1)(t - t_0)^k \phi_{\boldsymbol{w}}^{(k+1)}(\boldsymbol{w}, t_0). \tag{5.3.29}$$

由 (5.3.3) 式知道

$$\phi_{\boldsymbol{w}}^{(0)}(\boldsymbol{w}, t_0) = \phi_{\boldsymbol{w}}(\boldsymbol{w}, t_0, t_0) = \boldsymbol{f}(\boldsymbol{w}, t_0, t_0). \tag{5.3.30}$$

将 (5.3.28) 式代入 $\mathcal{A}\left(\widehat{\boldsymbol{w}}, \boldsymbol{w}, \dfrac{\partial \widehat{\boldsymbol{w}}}{\partial \boldsymbol{w}}, t, t_0\right)$, 并且将函数 $\mathcal{A}$ 在 $t = t_0$ 处展开, 可以得到

$$\begin{aligned}\mathcal{A}(\phi_{\boldsymbol{w}}, \boldsymbol{w}, \phi_{\boldsymbol{ww}}, t, t_0) = {} & \mathcal{A}(\boldsymbol{f}(\boldsymbol{w}, t_0, t_0), \boldsymbol{w}, \boldsymbol{f}_{\boldsymbol{w}}(\boldsymbol{w}, t_0, t_0), t_0, t_0) \\ & + \sum_{k=1}^{\infty}\frac{1}{k!}(t - t_0)^k \mathrm{D}_t^k \mathcal{A}(\phi_{\boldsymbol{w}}^{(0)}, \boldsymbol{w}, \phi_{\boldsymbol{ww}}^{(0)}, t_0, t_0).\end{aligned} \tag{5.3.31}$$

比较 (5.3.31) 式和 (5.3.29) 式, 再利用 (5.3.4) 式, 就可以得到 (5.3.26) 式和 (5.3.27) 式. □

上面依然是针对非自治情形的讨论, 自治和半自治情形的不同之处就是此处的函数 $\mathcal{A}$ 将被 Birkhoff 函数 $B(\boldsymbol{z}, t)$ 代替, 这使得生成函数 ϕ 的展开更容易. 定理 5.3.1 和定理 5.3.2 建立了 Birkhoff 系统相流和生成函数 ϕ 之间的明确关系. 基于这个结果, 理论上我们就可以构造任意阶的 Birkhoff 辛格式了.

定理 5.3.3 假设映射 $\boldsymbol{\alpha}$ 同命题 5.2.1 所定义, 函数 $\mathcal{A}$ 解析. 当步长 $\tau > 0$ 足够小时, 取

$$\psi_{\boldsymbol{w}}^{(m)}(\boldsymbol{w}, t_0 + \tau, t_0) = \sum_{i=0}^{m} \tau^i \phi_{\boldsymbol{w}}^{(i)}(\boldsymbol{w}, t_0), \quad m = 1, 2, \cdots, \tag{5.3.32}$$

这里 $\phi_{\boldsymbol{w}}^{(i)}$ 由方程 (5.3.25)—(5.3.27) 确定, 那么 $\psi_{\boldsymbol{w}}^{(m)}(\boldsymbol{w}, t_0 + \tau, t_0)$ 就定义了一个 m 阶精度的 $\boldsymbol{K}(\boldsymbol{z}, t)$-辛格式, 使得 $\boldsymbol{z} = \boldsymbol{z}^k \mapsto \boldsymbol{z}^{k+1} = \widehat{\boldsymbol{z}}$, 且

$$\boldsymbol{\alpha}_1(\boldsymbol{z}^{k+1}, \boldsymbol{z}^k, t_{k+1}, t_k) = \psi_{\boldsymbol{w}}^{(m)}(\boldsymbol{\alpha}_2(\boldsymbol{z}^{k+1}, \boldsymbol{z}^k, t_{k+1}, t_k), t_{k+1}, t_k). \tag{5.3.33}$$

证明 由于

$$\left|\boldsymbol{C}_{\boldsymbol{\alpha}}(\boldsymbol{z}, \boldsymbol{z}, t_0, t_0) + \boldsymbol{D}_{\boldsymbol{\alpha}}(\boldsymbol{z}, \boldsymbol{z}, t_0, t_0)\right| \neq 0, \tag{5.3.34}$$

因此由定理 5.2.2 知道 $|\boldsymbol{C}^{\boldsymbol{\alpha}}\boldsymbol{N} + \boldsymbol{D}^{\boldsymbol{\alpha}}| \neq 0$, 其中 $\boldsymbol{N}$ 是 $\phi_{\boldsymbol{w}}$ 的 Jacobi 矩阵: $\boldsymbol{N} = \phi_{\boldsymbol{ww}}(\boldsymbol{w}_0, t_0, t_0) = \psi_{\boldsymbol{ww}}^{(m)}(\boldsymbol{w}_0, t_0, t_0)$, $\boldsymbol{w}_0 = \boldsymbol{\alpha}_2(\boldsymbol{z}, \boldsymbol{z}, t_0, t_0)$. 所以, 当步长 τ 足够小时, 存在点 $\boldsymbol{w}_0$ 的某个邻域, 在该邻域内有

$$\left|\boldsymbol{C}^{\boldsymbol{\alpha}}\boldsymbol{N}^{(m)}(\boldsymbol{w}, t_0 + \tau, t_0) + \boldsymbol{D}^{\boldsymbol{\alpha}}\right| \neq 0, \tag{5.3.35}$$

这里的 $\boldsymbol{N}^{(m)}$ 是指

$$\boldsymbol{N}^{(m)}(\boldsymbol{w}, t_0 + \tau, t_0) = \psi_{\boldsymbol{ww}}^{(m)}(\boldsymbol{w}, t_0 + \tau, t_0). \tag{5.3.36}$$

由定理 5.2.2 知道, $\psi_{\boldsymbol{w}}^{(m)}(\boldsymbol{w}, t_0 + \tau, t_0)$ 定义了一个 $\boldsymbol{K}(\boldsymbol{z}, t)$-辛映射, 如方程 (5.2.35) 所表示, 所以说 (5.3.33) 式确定了一个 m 阶精度的 $\boldsymbol{K}(\boldsymbol{z}, t)$-辛格式. □

§5.4 带阻尼的振动方程的 Birkhoff 辛算法

在这一节中, 我们就一个具体的例子来说明如何应用上一节的理论来构造一个带有耗散项的非守恒系统的 Birkhoff 辛格式[SQ04]. 考虑带阻尼的振动方程

$$\ddot{q} + \nu\dot{q} + q = 0. \tag{5.4.1}$$

在这个方程中, 有耗散项 $\nu\dot{q}$, 它的总能量 $H=\dfrac{1}{2}(\dot{q}^2+q^2)-\nu\dot{q}^2$ 显然不守恒. 这是一个简单的 Newton 方程, 它不满足对称条件, 所以要考虑引进新变量. 令 $p=\dot{q}$, 那么系统 (5.4.1) 就可以写成如下形式:

$$\dot{p}=-\nu p-q,\quad \dot{q}=p, \tag{5.4.2}$$

进而写成 Birkhoff 形式

$$\begin{pmatrix} 0 & -\mathrm{e}^{\nu t} \\ \mathrm{e}^{\nu t} & 0 \end{pmatrix}\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix}=\begin{pmatrix} \nu\mathrm{e}^{\nu t}p+\mathrm{e}^{\nu t}q \\ \mathrm{e}^{\nu t}p \end{pmatrix}. \tag{5.4.3}$$

它的 $\boldsymbol{K}(\boldsymbol{z},t)$-辛结构和 Birkhoff 函数分别是

$$\begin{aligned} &\boldsymbol{K}=\begin{pmatrix} 0 & -\mathrm{e}^{\nu t} \\ \mathrm{e}^{\nu t} & 0 \end{pmatrix},\quad \boldsymbol{K}^{-1}=\begin{pmatrix} 0 & \mathrm{e}^{-\nu t} \\ -\mathrm{e}^{-\nu t} & 0 \end{pmatrix}, \\ &\boldsymbol{F}=\begin{pmatrix} \dfrac{1}{2}\mathrm{e}^{\nu t}p \\ -\dfrac{1}{2}\mathrm{e}^{\nu t}q \end{pmatrix},\quad B=\frac{1}{2}\mathrm{e}^{\nu t}(q^2+\nu qp+p^2). \end{aligned} \tag{5.4.4}$$

5.4.1 Birkhoff 辛格式

令 $(\widehat{\boldsymbol{z}},\boldsymbol{z})=(\widehat{q},\widehat{p},q,p),(\widehat{\boldsymbol{w}},\boldsymbol{w})=(\widehat{Q},\widehat{P},Q,P)$. 取一个简单的 $\boldsymbol{\alpha}$ 变换:

$$\begin{aligned} &\widehat{Q}=\mathrm{e}^{\nu t}\widehat{p}-\mathrm{e}^{\nu t_0}p,\quad \widehat{P}=\widehat{q}-q, \\ &Q=\frac{1}{2}(\widehat{q}+q),\quad P=-\frac{1}{2}(\mathrm{e}^{\nu t}\widehat{p}+\mathrm{e}^{\nu t_0}p), \end{aligned} \tag{5.4.5}$$

它的 Jacobi 矩阵是

$$\boldsymbol{\alpha}_*=\begin{pmatrix} 0 & \mathrm{e}^{\nu t} & 0 & -\mathrm{e}^{\nu t_0} \\ 1 & 0 & -1 & 0 \\ \dfrac{1}{2} & 0 & \dfrac{1}{2} & 0 \\ 0 & -\dfrac{1}{2}\mathrm{e}^{\nu t} & 0 & -\dfrac{1}{2}\mathrm{e}^{\nu t_0} \end{pmatrix}. \tag{5.4.6}$$

$\boldsymbol{\alpha}$ 的逆变换是

$$\begin{aligned}&\widehat{q}=\frac{1}{2}\widehat{P}+Q,\quad \widehat{p}=\frac{1}{2}\mathrm{e}^{-\nu t}\widehat{Q}-\mathrm{e}^{-\nu t}P,\\&q=-\frac{1}{2}\widehat{P}+Q,\quad p=-\frac{1}{2}\mathrm{e}^{\nu t_0}\widehat{Q}-\mathrm{e}^{-\nu t_0}P,\end{aligned}\tag{5.4.7}$$

它的 Jacobi 矩阵是

$$\boldsymbol{\alpha}_*^{-1}=\begin{pmatrix}0&\dfrac{1}{2}&1&0\\\dfrac{1}{2}\mathrm{e}^{-\nu t}&0&0&-\mathrm{e}^{-\nu t}\\0&-\dfrac{1}{2}&1&0\\-\dfrac{1}{2}\mathrm{e}^{-\nu t_0}&0&0&-\mathrm{e}^{-\nu t_0}\end{pmatrix}.\tag{5.4.8}$$

由变换 $\boldsymbol{\alpha}$ 和它的逆变换得到

$$\begin{aligned}&\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}=\begin{pmatrix}\nu\mathrm{e}^{\nu t}\widehat{p}+\mathrm{e}^{\nu t}\dot{\widehat{p}}\\\dot{\widehat{q}}\end{pmatrix}=\begin{pmatrix}-\mathrm{e}^{\nu t}\widehat{q}\\\widehat{p}\end{pmatrix}=\begin{pmatrix}-\dfrac{1}{2}\mathrm{e}^{\nu t}\widehat{P}-\mathrm{e}^{\nu t}Q\\\dfrac{1}{2}\mathrm{e}^{-\nu t}\widehat{Q}-\mathrm{e}^{-\nu t}P\end{pmatrix},\\&\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}=\begin{pmatrix}\dfrac{1}{4}\mathrm{e}^{-\nu t}\widehat{Q}-\dfrac{1}{2}\mathrm{e}^{-\nu t}P\\\dfrac{1}{4}\mathrm{e}^{\nu t}\widehat{P}+\dfrac{1}{2}\mathrm{e}^{\nu t}Q\end{pmatrix}.\end{aligned}\tag{5.4.9}$$

经过很简单的计算, 就可以得到

$$\begin{aligned}&\phi_{\boldsymbol{w}}^{(0)}=\begin{pmatrix}\widehat{Q}\\\widehat{P}\end{pmatrix}\bigg|_{t=t_0}=\begin{pmatrix}0\\0\end{pmatrix},\\&\phi_{\boldsymbol{w}}^{(1)}=\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\bigg|_{t=t_0}-\phi_{\boldsymbol{ww}}^{(0)}\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\bigg|_{t=t_0}=\begin{pmatrix}-\mathrm{e}^{\nu t_0}Q\\-\mathrm{e}^{-\nu t_0}P\end{pmatrix}.\end{aligned}\tag{5.4.10}$$

令

$$\widehat{\boldsymbol{w}}=\phi_{\boldsymbol{w}}^{(0)}+\phi_{\boldsymbol{w}}^{(1)}\tau,\tag{5.4.11}$$

这样就得到一个 1 阶精度的 Birkhoff 辛格式:

$$\begin{aligned}&\frac{q_{k+1}-q_k}{\tau}=\mathrm{e}^{-\nu t_k}\frac{\mathrm{e}^{\nu t_{k+1}}p_{k+1}+\mathrm{e}^{\nu t_k}p_k}{2},\\&\frac{\mathrm{e}^{\nu t_{k+1}}p_{k+1}-\mathrm{e}^{\nu t_k}p_k}{\tau}=-\mathrm{e}^{\nu t_k}\frac{q_{k+1}+q_k}{2}.\end{aligned}\tag{5.4.12}$$

格式 (5.4.12) 可整理成如下紧凑形式:

$$\begin{pmatrix} q_{k+1} \\ p_{k+1} \end{pmatrix} = \begin{pmatrix} \dfrac{4-\tau^2}{4+\tau^2} & \dfrac{4\tau}{4+\tau^2} \\ -\dfrac{4\tau}{4+\tau^2}\mathrm{e}^{-\nu\tau} & \dfrac{4-\tau^2}{4+\tau^2}\mathrm{e}^{-\nu\tau} \end{pmatrix} \begin{pmatrix} q_k \\ p_k \end{pmatrix}. \tag{5.4.13}$$

令

$$\boldsymbol{A} = \begin{pmatrix} \dfrac{4-\tau^2}{4+\tau^2} & \dfrac{4\tau}{4+\tau^2} \\ -\dfrac{4\tau}{4+\tau^2}\mathrm{e}^{-\nu\tau} & \dfrac{4-\tau^2}{4+\tau^2}\mathrm{e}^{-\nu\tau} \end{pmatrix},$$

易证

$$\boldsymbol{A}^{\mathrm{T}} \begin{pmatrix} 0 & -\mathrm{e}^{\nu t_{k+1}} \\ \mathrm{e}^{\nu t_{k+1}} & 0 \end{pmatrix} \boldsymbol{A} = \begin{pmatrix} 0 & -\mathrm{e}^{\nu t_k} \\ \mathrm{e}^{\nu t_k} & 0 \end{pmatrix}, \tag{5.4.14}$$

所以格式 (5.4.12) 即格式 (5.4.13) 是 Birkhoff 辛格式.

另外, 对格式 (5.4.12) 做修正, 可得到一个 2 阶精度的 Birkhoff 辛格式:

$$\begin{aligned} \frac{q_{k+1}-q_k}{\tau} &= \mathrm{e}^{-\nu t_k+\tau/2}\frac{\mathrm{e}^{\nu t_{k+1}}p_{k+1}+\mathrm{e}^{\nu t_k}p_k}{2}, \\ \frac{\mathrm{e}^{\nu t_{k+1}}p_{k+1}-\mathrm{e}^{\nu t_k}p_k}{\tau} &= -\mathrm{e}^{\nu t_k+\tau/2}\frac{q_{k+1}+q_k}{2}. \end{aligned} \tag{5.4.15}$$

继续利用迭代公式 (5.3.27) 可求得

$$\phi_{\boldsymbol{w}}^{(2)} = \begin{pmatrix} -\dfrac{\nu}{2}\mathrm{e}^{\nu t_0}Q \\ \dfrac{\nu}{2}\mathrm{e}^{-\nu t_0}P \end{pmatrix}. \tag{5.4.16}$$

令 $\widehat{\boldsymbol{w}} = \phi_{\boldsymbol{w}}^{(0)} + \phi_{\boldsymbol{w}}^{(1)}\tau + \phi_{\boldsymbol{w}}^{(2)}\tau^2$, 得到系统 (5.4.3) 的又一个 2 阶精度的 Birkhoff 辛格式:

$$\begin{pmatrix} q_{k+1} \\ p_{k+1} \end{pmatrix} = \begin{pmatrix} \dfrac{16-ab}{16+ab} & \dfrac{8a}{16+ab} \\ -\dfrac{8b}{16+ab}\mathrm{e}^{-\nu\tau} & \dfrac{16-ab}{16+ab}\mathrm{e}^{-\nu\tau} \end{pmatrix} \begin{pmatrix} q_k \\ p_k \end{pmatrix}, \tag{5.4.17}$$

其中

$$a = 2\tau - \nu\tau^2, \quad b = 2\tau + \nu\tau^2. \tag{5.4.18}$$

格式 (5.4.17) 也是 Birkhoff 辛格式.

再利用 Euler 中点格式直接离散系统 (5.4.3), 得到下面的格式:

$$\begin{pmatrix} q_{k+1} \\ p_{k+1} \end{pmatrix} = \begin{pmatrix} \dfrac{-\tau^2+2\nu\tau+4}{\tau^2+2\nu\tau+4} & \dfrac{4\tau}{\tau^2+2\nu\tau+4} \\ \dfrac{-4\tau}{\tau^2+2\nu\tau+4} & \dfrac{-\tau^2-2\nu\tau+4}{\tau^2+2\nu\tau+4} \end{pmatrix} \begin{pmatrix} q_k \\ p_k \end{pmatrix}. \tag{5.4.19}$$

可以证明这不是一个 Birkhoff 辛格式.

若令 (5.2.11) 式中的变换 $\boldsymbol{\alpha}$ 取为

$$\begin{aligned} &\widehat{Q} = \mathrm{e}^{\nu t/2}\widehat{p} - \mathrm{e}^{\nu t_0/2}p, \quad \widehat{P} = -\mathrm{e}^{\nu t/2}\widehat{q} + \mathrm{e}^{\nu t_0/2}q, \\ &Q = \frac{1}{2}(\mathrm{e}^{\nu t/2}\widehat{q} + \mathrm{e}^{\nu t_0/2}q), \quad P = \frac{1}{2}(\mathrm{e}^{\nu t/2}\widehat{p} + \mathrm{e}^{\nu t_0/2}p), \end{aligned} \tag{5.4.20}$$

则 $\boldsymbol{\alpha}$ 的 Jacobi 矩阵为

$$\boldsymbol{\alpha}_* = \begin{pmatrix} 0 & \mathrm{e}^{\nu t/2} & 0 & -\mathrm{e}^{\nu t_0/2} \\ -\mathrm{e}^{\nu t/2} & 0 & \mathrm{e}^{\nu t_0/2} & 0 \\ \dfrac{1}{2}\mathrm{e}^{\nu t/2} & 0 & \dfrac{1}{2}\mathrm{e}^{\nu t_0/2} & 0 \\ 0 & \dfrac{1}{2}\mathrm{e}^{\nu t/2} & 0 & \dfrac{1}{2}\mathrm{e}^{\nu t_0/2} \end{pmatrix}, \tag{5.4.21}$$

其逆矩阵为

$$\boldsymbol{\alpha}_*^{-1} = \begin{pmatrix} 0 & -\dfrac{1}{2}\mathrm{e}^{-\nu t/2} & \mathrm{e}^{-\nu t/2} & 0 \\ \dfrac{1}{2}\mathrm{e}^{-\nu t/2} & 0 & 0 & \mathrm{e}^{-\nu t/2} \\ 0 & \dfrac{1}{2}\mathrm{e}^{-\nu t_0/2} & \mathrm{e}^{-\nu t_0/2} & 0 \\ -\dfrac{1}{2}\mathrm{e}^{-\nu t_0/2} & 0 & 0 & \mathrm{e}^{-\nu t_0/2} \end{pmatrix}. \tag{5.4.22}$$

取 $m=1$ 次截断, 得到 2 阶精度的对称格式

$$\widehat{\boldsymbol{w}} = \phi_{\boldsymbol{w}}^{(0)} + \phi_{\boldsymbol{w}}^{(1)}\tau,$$

其中

$$\phi_{\boldsymbol{w}}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \phi_{\boldsymbol{w}}^{(1)} = \left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t} - \frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right)\Big|_{t=t_0} = \begin{pmatrix} -\dfrac{\nu}{2}P - Q \\ -\dfrac{\nu}{2}Q - P \end{pmatrix}.$$

由此可得到 2 阶精度的对称辛格式

$$
\begin{aligned}
\frac{e^{\nu t_{k+1}/2}q_{k+1}-e^{\nu t_k/2}q_k}{\tau}&=\frac{e^{\nu t_{k+1}/2}p_{k+1}+e^{\nu t_k/2}p_k}{2}\\
&\quad+\nu\frac{e^{\nu t_{k+1}/2}q_{k+1}+e^{\nu t_k/2}q_k}{4},\\
\frac{e^{\nu t_{k+1}/2}p_{k+1}-e^{\nu t_k/2}p_k}{\tau}&=-\frac{e^{\nu t_{k+1}/2}q_{k+1}+e^{\nu t_k/2}q_k}{2}\\
&\quad-\nu\frac{e^{\nu t_{k+1}/2}p_{k+1}+e^{\nu t_k/2}p_k}{4},
\end{aligned}
\tag{5.4.23}
$$

其紧凑形式为

$$
\begin{pmatrix} q_{k+1}\\ p_{k+1}\end{pmatrix}=\frac{e^{-\nu\tau/2}}{\Delta}\begin{pmatrix} w_1 & -16\tau\\ 16\tau & w_2\end{pmatrix}\begin{pmatrix} q_k\\ p_k\end{pmatrix},
\tag{5.4.24}
$$

其中

$$
\begin{aligned}
\Delta&=\nu^2\tau^2-4\tau^2-16,\\
w_1&=-16-8\nu\tau-\nu^2\tau^2+4\tau^2,\\
w_2&=-16+8\nu\tau-\nu^2\tau^2+4\tau^2.
\end{aligned}
$$

详细推导可参见本章附录.

我们把格式 (5.4.24) 中的矩阵 $\dfrac{e^{-\nu\tau/2}}{\Delta}\begin{pmatrix} w_1 & -16\tau\\ 16\tau & w_2\end{pmatrix}$ 记为 $\boldsymbol{M}(\tau)$, 则利用组合格式理论[Yos90,QZ92]可得到一个组合格式

$$
\begin{pmatrix} q_{k+1}\\ p_{k+1}\end{pmatrix}=\boldsymbol{M}(c_1\tau)\boldsymbol{M}(c_2\tau)\boldsymbol{M}(c_1\tau)\begin{pmatrix} q_k\\ p_k\end{pmatrix},
\tag{5.4.25}
$$

其中

$$
c_1=\frac{1}{2-2^{1/3}},\quad c_2=\frac{-2^{1/3}}{2-2^{1/3}}.
$$

这个格式也是 Birkhoff 辛格式.

若变换 (5.4.20) 取 $m=2$ 阶截断, 可计算出

$$\phi_{\boldsymbol{w}}^{(2)}=0.$$

这也说明了 1 阶截断的格式 (5.4.23) 为何具有 2 阶精度. 今取 $m=3$, 则

$$\begin{aligned}\phi_{\boldsymbol{w}}^{(3)}=&\frac{1}{3!}\frac{\partial}{\partial t}\left[\frac{\partial}{\partial t}\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right)+\frac{\partial}{\partial\widehat{\boldsymbol{w}}}\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right)\frac{\partial\widehat{\boldsymbol{w}}}{\partial t}-\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\frac{\partial}{\partial\widehat{\boldsymbol{w}}}\left(\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right)\frac{\partial\widehat{\boldsymbol{w}}}{\partial t}\right.\\&\left.-\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\frac{\partial}{\partial t}\left(\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right)-\frac{\partial}{\partial t}\left(\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\right)\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right].\end{aligned}$$

对于方程 $\ddot{q}+\nu\dot{q}+q=0$, 3 阶导数 $\phi_{\boldsymbol{w}}^{(3)}$ 在 $t=t_0$ 取值时, 上式仅剩下一项, 即

$$-\frac{\partial}{\partial t}\left(\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\right)\frac{\partial}{\partial\widehat{\boldsymbol{w}}}\left(\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right)\frac{\partial\widehat{\boldsymbol{w}}}{\partial t},$$

故

$$\begin{aligned}\left.\phi_{\boldsymbol{w}}^{(3)}\right|_{t=t_0}&=-\frac{1}{6}\begin{pmatrix}-1&-\dfrac{\nu}{2}\\-\dfrac{\nu}{2}&-1\end{pmatrix}\begin{pmatrix}\dfrac{1}{4}&-\dfrac{\nu}{8}\\-\dfrac{\nu}{8}&\dfrac{1}{4}\end{pmatrix}\begin{pmatrix}-\dfrac{\nu}{2}P-Q\\\dfrac{-\nu}{2}Q-P\end{pmatrix}\\&=-\frac{1}{6}\begin{pmatrix}\left(-\dfrac{1}{4}+\dfrac{1}{16}\nu^2\right)\left(Q+\dfrac{\nu P}{2}\right)\\\left(-\dfrac{1}{4}+\dfrac{1}{16}\nu^2\right)\left(\dfrac{\nu Q}{2}+P\right)\end{pmatrix}\\&=\frac{1}{48}\begin{pmatrix}\left(\dfrac{\nu^2}{2}-2\right)Q+\left(\dfrac{\nu^2}{2}-2\right)\dfrac{\nu}{2}P\\\left(\dfrac{\nu^2}{2}-2\right)P+\left(\dfrac{\nu^2}{2}-2\right)\dfrac{\nu}{2}Q\end{pmatrix}.\end{aligned}$$

由此可得 4 阶精度的对称格式

$$\widehat{\boldsymbol{w}}=\phi_{\boldsymbol{w}}^{(1)}\tau+\phi_{\boldsymbol{w}}^{(3)}\tau^3,$$

即

$$
\begin{aligned}
&\frac{\mathrm{e}^{\nu t_{k+1}/2}q_{k+1}-\mathrm{e}^{\nu t_k/2}q_k}{\tau}\\
&=\frac{\mathrm{e}^{\nu t_{k+1}/2}p_{k+1}+\mathrm{e}^{\nu t_k/2}p_k}{2}+\nu\frac{\mathrm{e}^{\nu t_{k+1}/2}q_{k+1}-\mathrm{e}^{\nu t_k/2}q_k}{4}\\
&\quad-\tau^2\frac{1}{24\times 4}\left[\left(\frac{\nu^2}{2}-2\right)(\mathrm{e}^{\nu t_{k+1}/2}p_{k+1}+\mathrm{e}^{\nu t_k/2}p_k)\right.\\
&\quad\left.+\left(\frac{\nu^2}{2}-2\right)\frac{\nu}{2}(\mathrm{e}^{\nu t_{k+1}/2}q_{k+1}+\mathrm{e}^{\nu t_k/2}q_k)\right],\\
&\frac{\mathrm{e}^{\nu t_{k+1}/2}p_{k+1}-\mathrm{e}^{\nu t_k/2}p_k}{\tau}\\
&=-\frac{\mathrm{e}^{\nu t_{k+1}/2}q_{k+1}+\mathrm{e}^{\nu t_k/2}q_k}{2}-\nu\frac{\mathrm{e}^{\nu t_{k+1}/2}p_{k+1}+\mathrm{e}^{\nu t_k/2}p_k}{4}\\
&\quad+\tau^2\frac{1}{24\times 4}\left[\left(\frac{\nu^2}{2}-2\right)(\mathrm{e}^{\nu t_{k+1}/2}q_{k+1}+\mathrm{e}^{\nu t_k/2}q_k)\right.\\
&\quad\left.+\left(\frac{\nu^2}{2}-2\right)\frac{\nu}{2}(\mathrm{e}^{\nu t_{k+1}/2}p_{k+1}+\mathrm{e}^{\nu t_k/2}p_k)\right].
\end{aligned}
\tag{5.4.26}
$$

注 5.4.1 格式 (5.4.12), (5.4.15), (5.4.17), (5.4.23), (5.4.25) 和 (5.4.26) 对任何 $\tau>0$ 都是 $\boldsymbol{K}(\boldsymbol{z},t)$-辛的, 即对 $k>0$ 满足

$$
\mathrm{e}^{\nu t_{k+1}}\mathrm{d}q_{k+1}\wedge\mathrm{d}p_{k+1}=\mathrm{e}^{\nu t_k}\mathrm{d}q_k\wedge\mathrm{d}p_k. \tag{5.4.27}
$$

上面的生成函数方法可以推广到如下一般常微分方程系统:

$$
\begin{aligned}
&\dot{p}+\beta'(t)p+V(q,t)=0,\\
&\dot{q}-G(p,t)=0,
\end{aligned}
\tag{5.4.28}
$$

其中 $\beta(t),V(q,t)$ 和 $G(p,t)$ 都是光滑函数. 显然, 当 $\beta'(t)=\nu,V(q,t)=q$, $G(p,t)=p$ 时, 系统 (5.4.28) 就是系统 (5.4.2).

5.4.2 数值实验

本节将给出线性耗散系统 (5.4.1) 基于 Birkhoff 形式 (5.4.3) 的数值模拟结果. 对 $\boldsymbol{K}(\boldsymbol{z},t)$-辛格式 (5.4.12), (5.4.23), (5.4.25) 和非辛

Euler 中点格式 (5.4.19) (此格式与 Marsden 等人[MW01] 利用 Lagrange-d'Alembert 变分原理对受外力作用系统作离散得到的格式是一致的) 进行数值比较, 同时作比较的还有显式 Euler 格式. 结果表明, Birkhoff 辛格式都能较好地模拟原系统, 且能更好地反映系统能量的耗散过程.

数值实验初值选择为 $q(0)=1, p(0)=\dot{q}(0)=-1$, 时间从 0 计算到 25. 数值实验表明, 不同初值的选取对于不同格式性态差别影响很小. 耗散因子 ν 分别取为 0.6, 1.3, 1.9, 2.8, 实际误差 err = |近似解 − 精确解|, 利用不同步长 τ 计算最大误差 $\mathrm{err}^{\max}$. 图形采用对数尺度, 利用不同时间步长 $\tau = 0.01, 0.02, 0.05, 0.1, 0.2, 0.5$ 来表现各个格式的性态.

首先对 $\boldsymbol{K}(z,t)$-辛格式 (5.4.12) 与 1 阶精度的显式 Euler 格式作比较: 对于较小的 ν, 比如 $0 < \nu \leqslant 1.3$, 格式 (5.4.12) 比显式 Euler 格式好; 当 $\nu > 1.3$ 时, 显式 Euler 格式优于格式 (5.4.12). 再将 2 阶精度的 Birkhoff 辛格式 (5.4.23) 和 Euler 中点格式 (5.4.19) 比较: 当 $0 < \nu \leqslant 1.3$ 时, 两格式性态差不多, 格式 (5.4.23) 略微优于格式 (5.4.19); 当 $1.3 < \nu < 2.5$ 时, 格式 (5.4.23) 明显优于格式 (5.4.19). 组合格式 (5.4.25) 则始终大大优于其他格式. 这些格式的比较参见图 5.4.1—5.4.4.

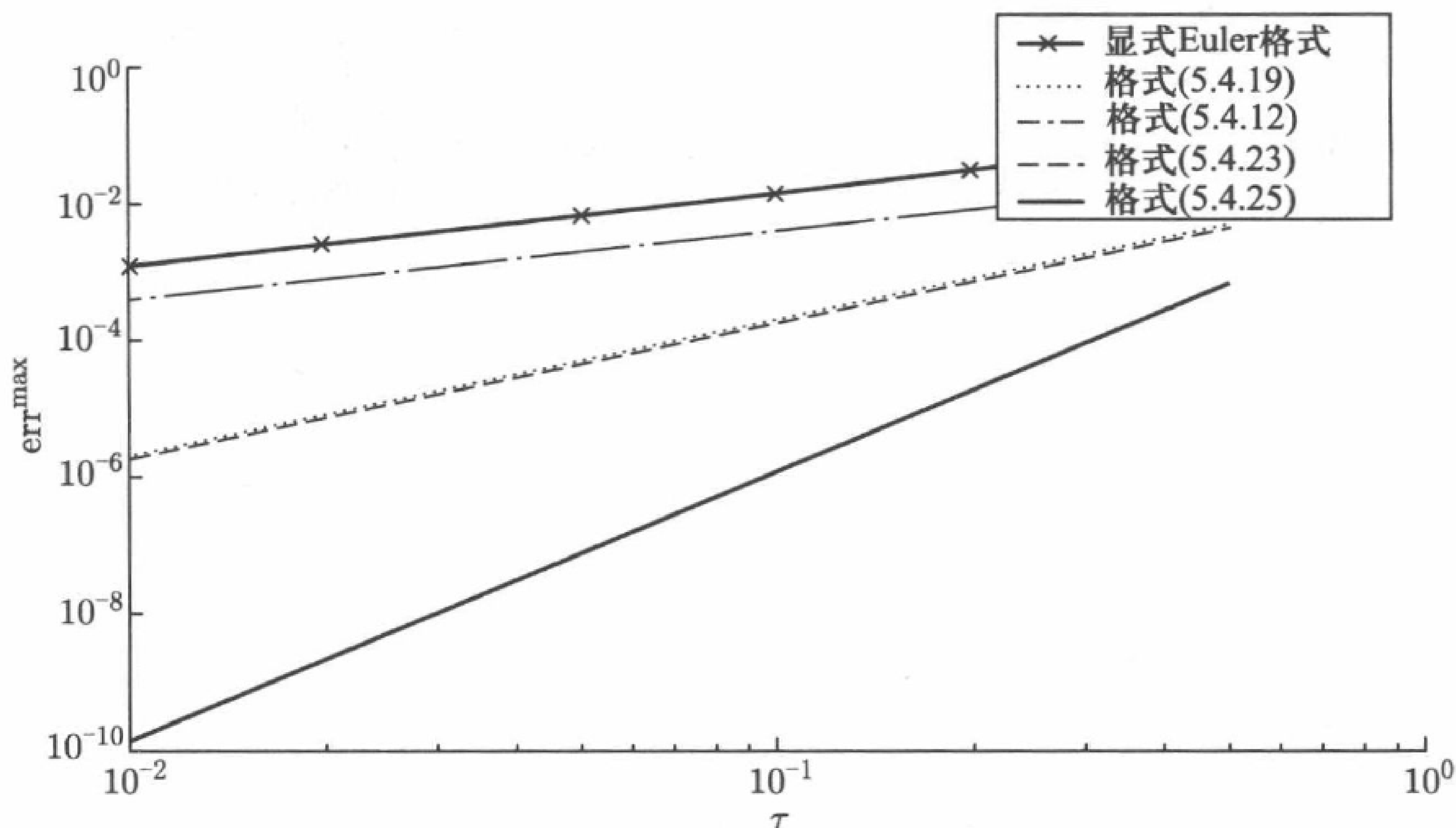


图 5.4.1 $\nu = 0.6$ 时不同差分格式的误差比较

同时利用以上几个格式计算能量函数 $H = \dfrac{1}{2}(\dot{q}^2 + q^2) - \nu p^2$, 对比

各个格式计算的能量最大误差 $\mathrm{err}_H^{\max}$, 结果见图 5.4.5—5.4.8.

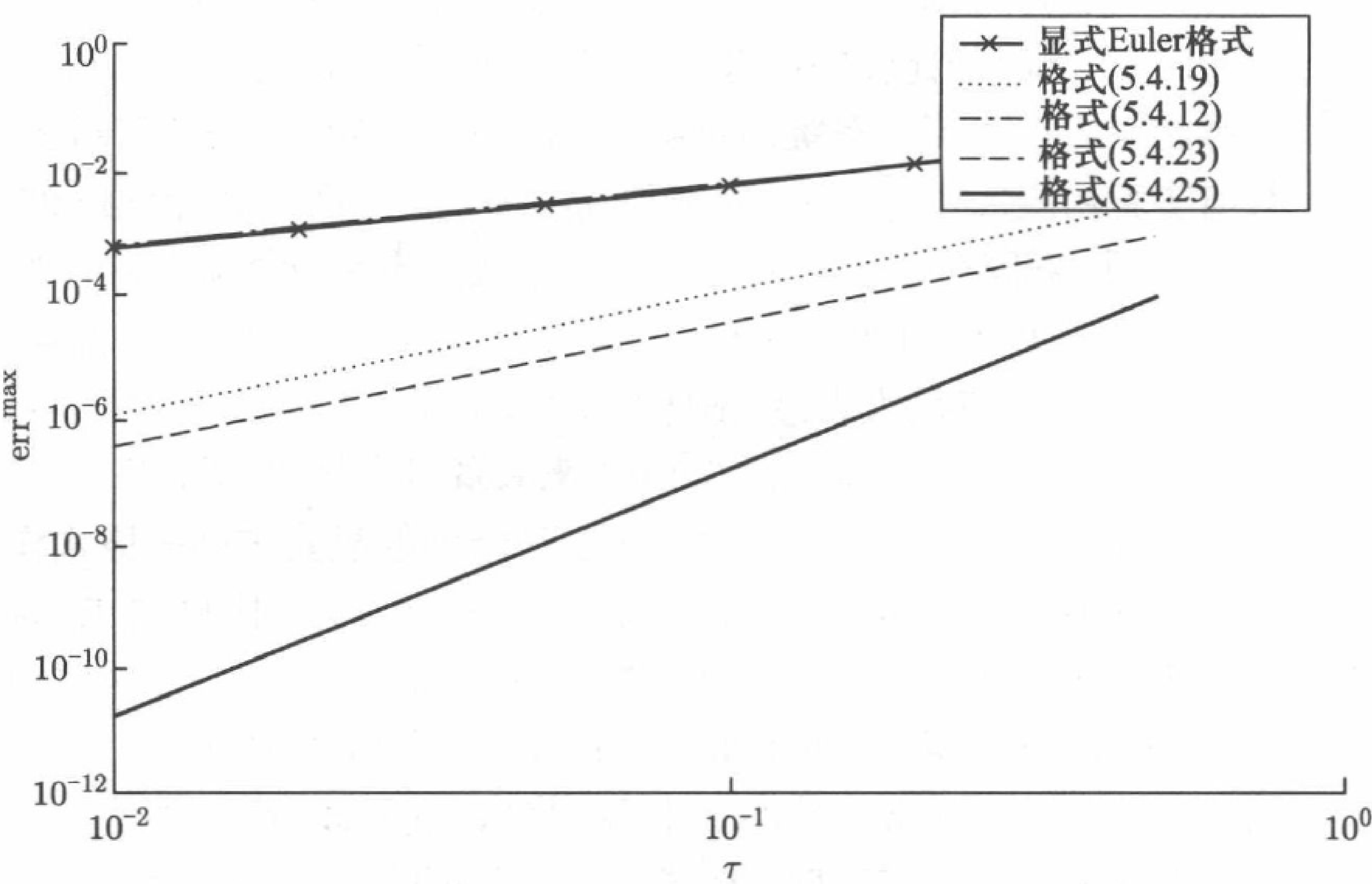


图 5.4.2　$\nu = 1.3$ 时不同差分格式的误差比较

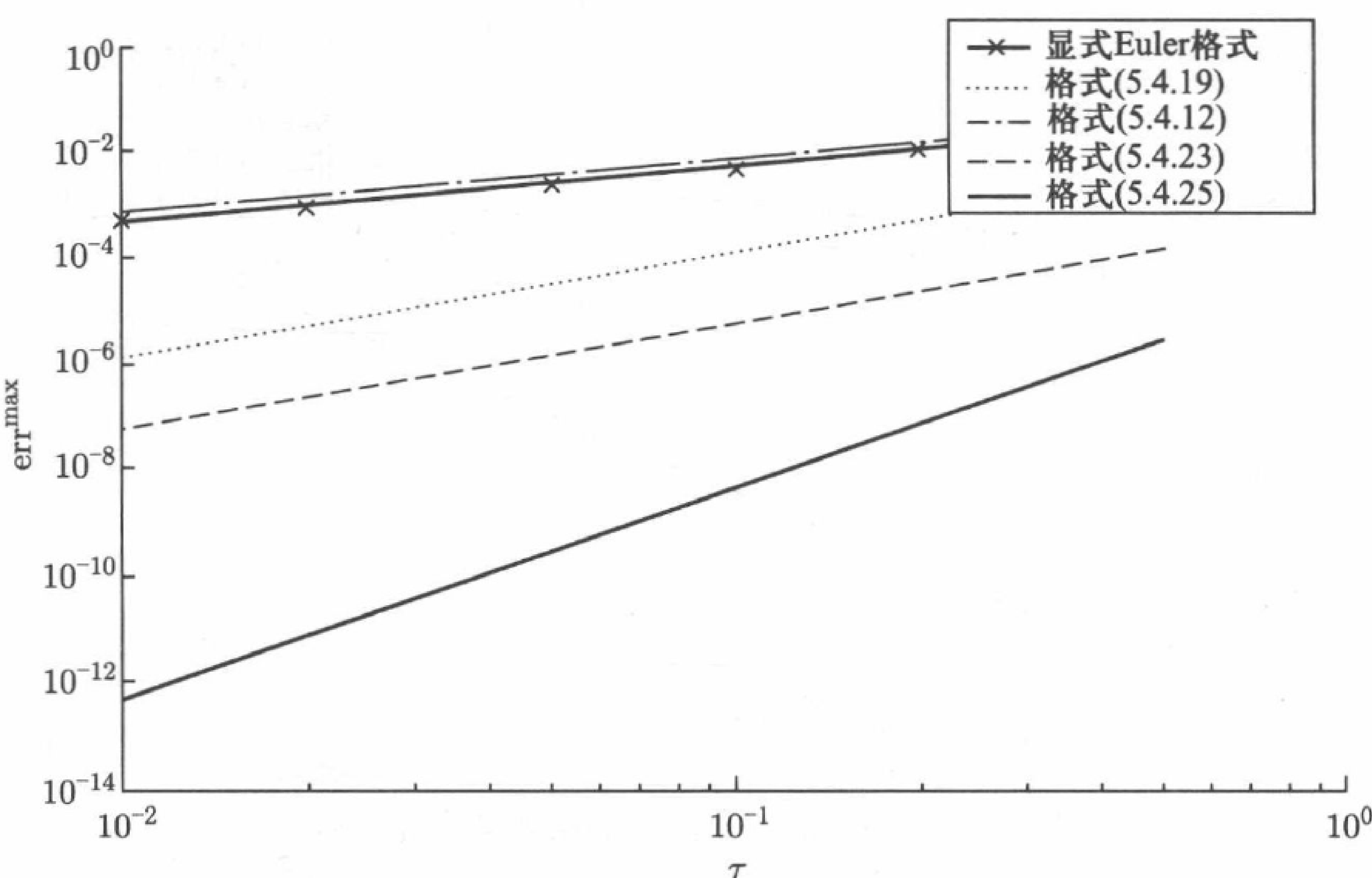


图 5.4.3　$\nu = 1.9$ 时不同差分格式的误差比较

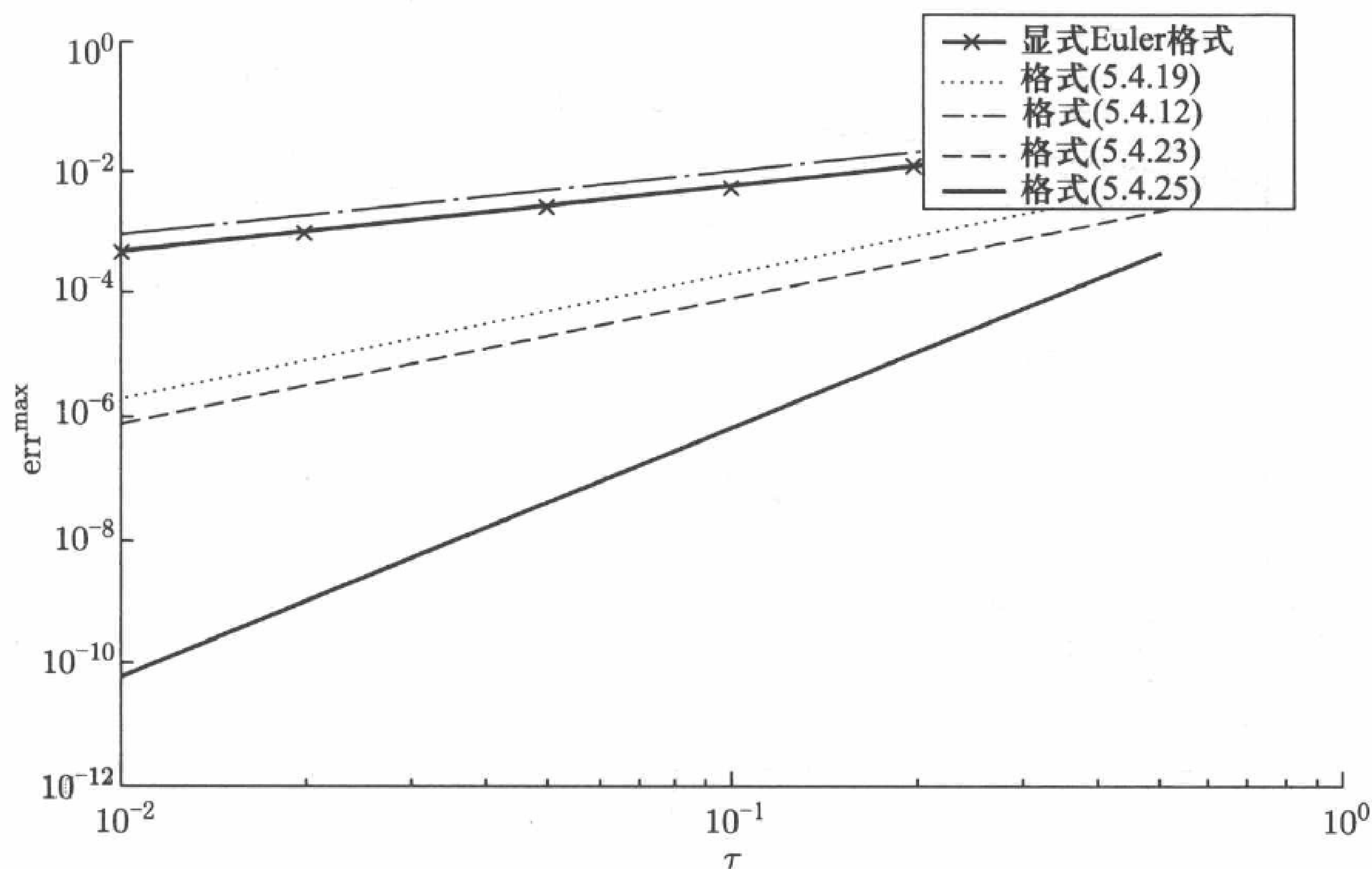


图 5.4.4 $\nu = 2.8$ 时不同差分格式的误差比较

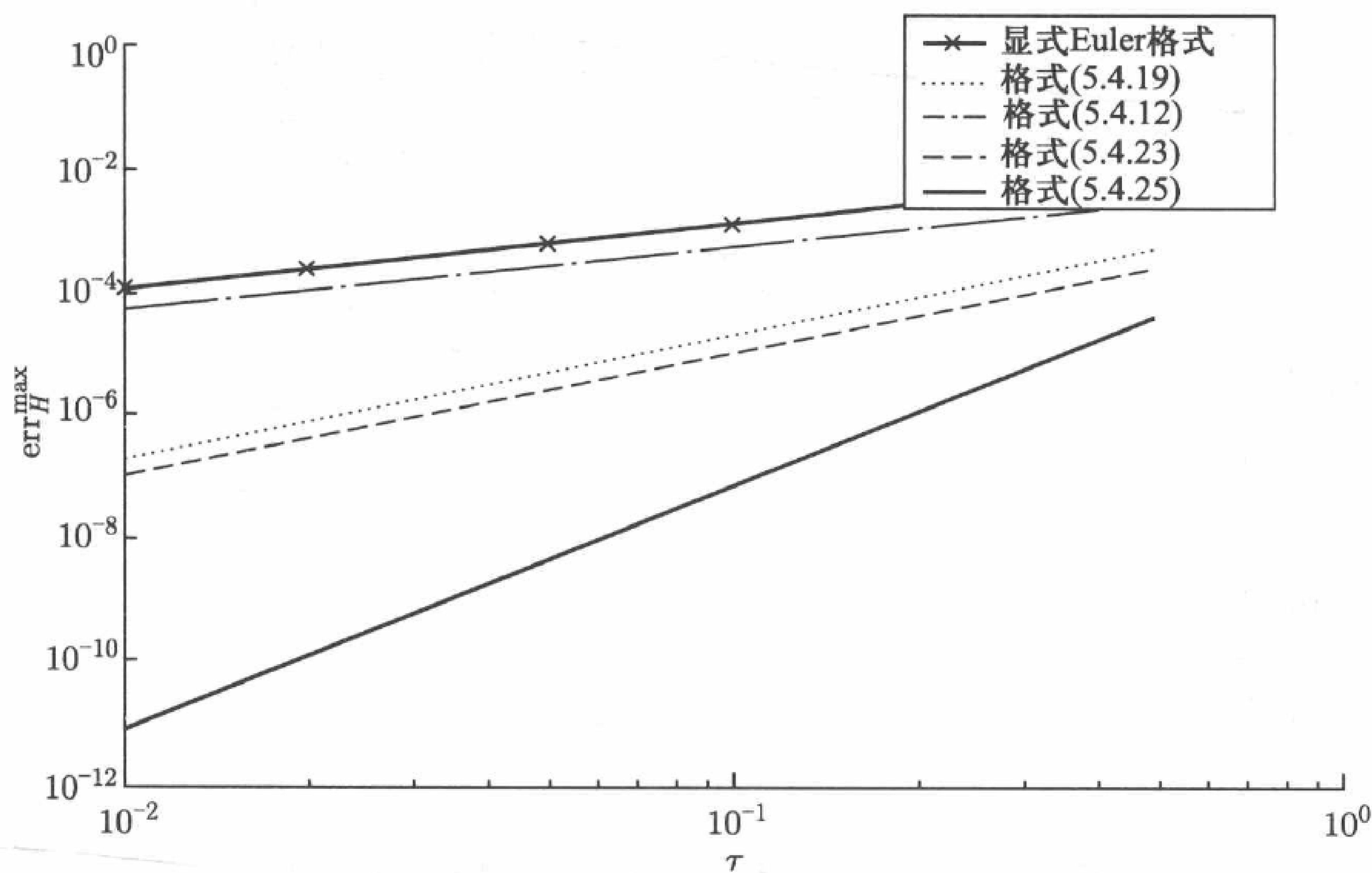


图 5.4.5 $\nu = 0.6$ 时不同差分格式的能量误差比较

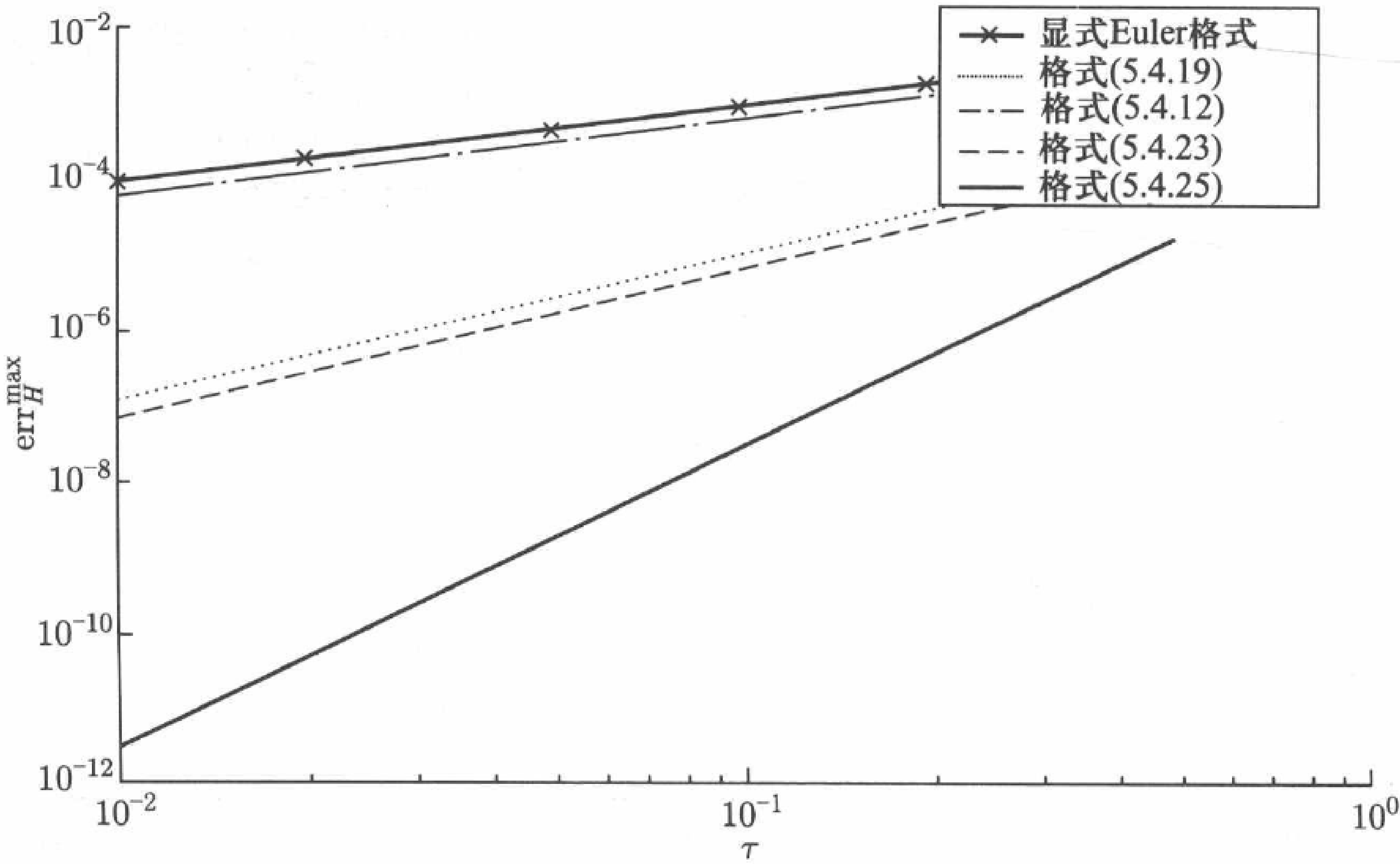


图 5.4.6　$\nu = 1.3$ 时不同差分格式的能量误差比较

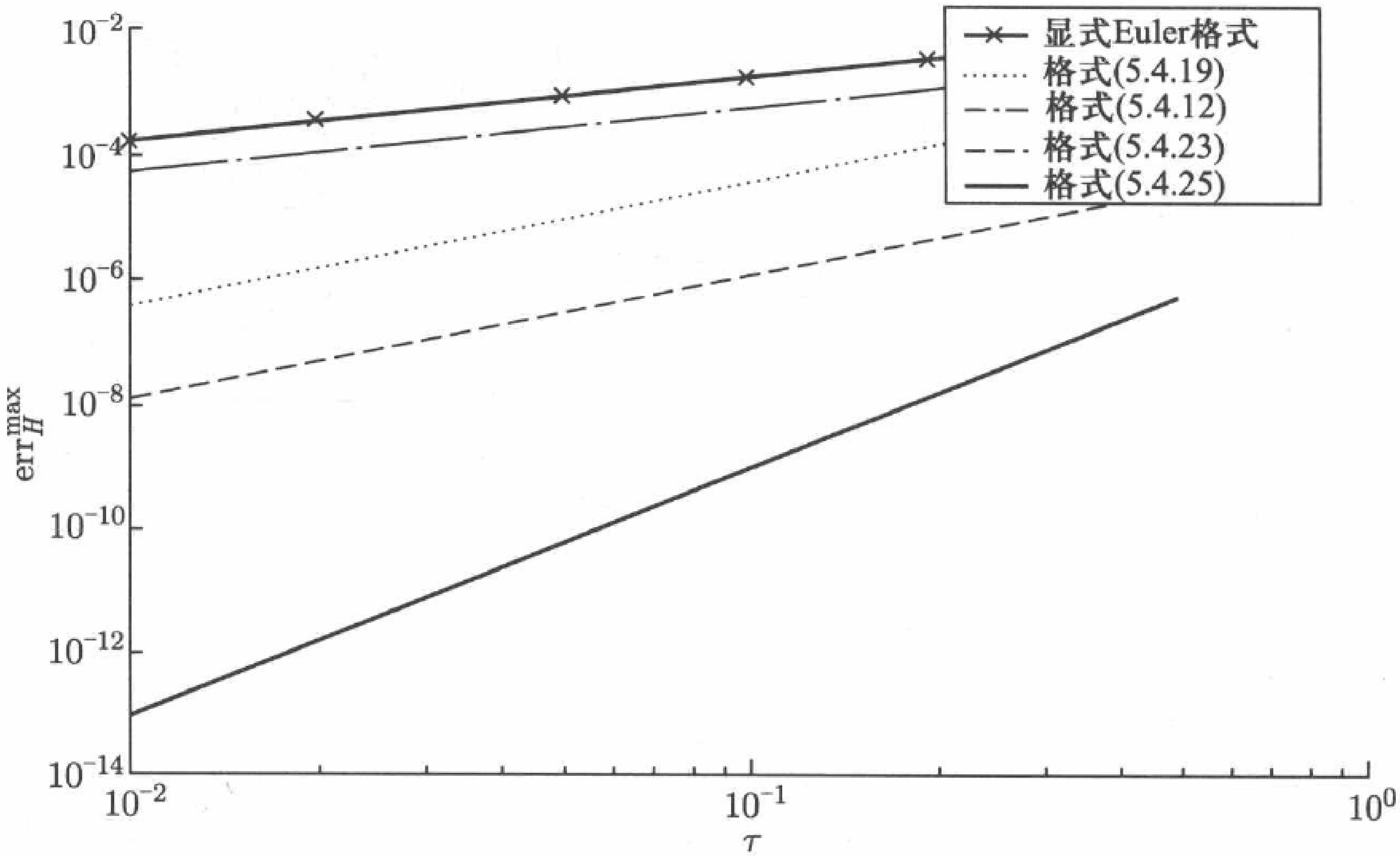


图 5.4.7　$\nu = 1.9$ 时不同差分格式的能量误差比较

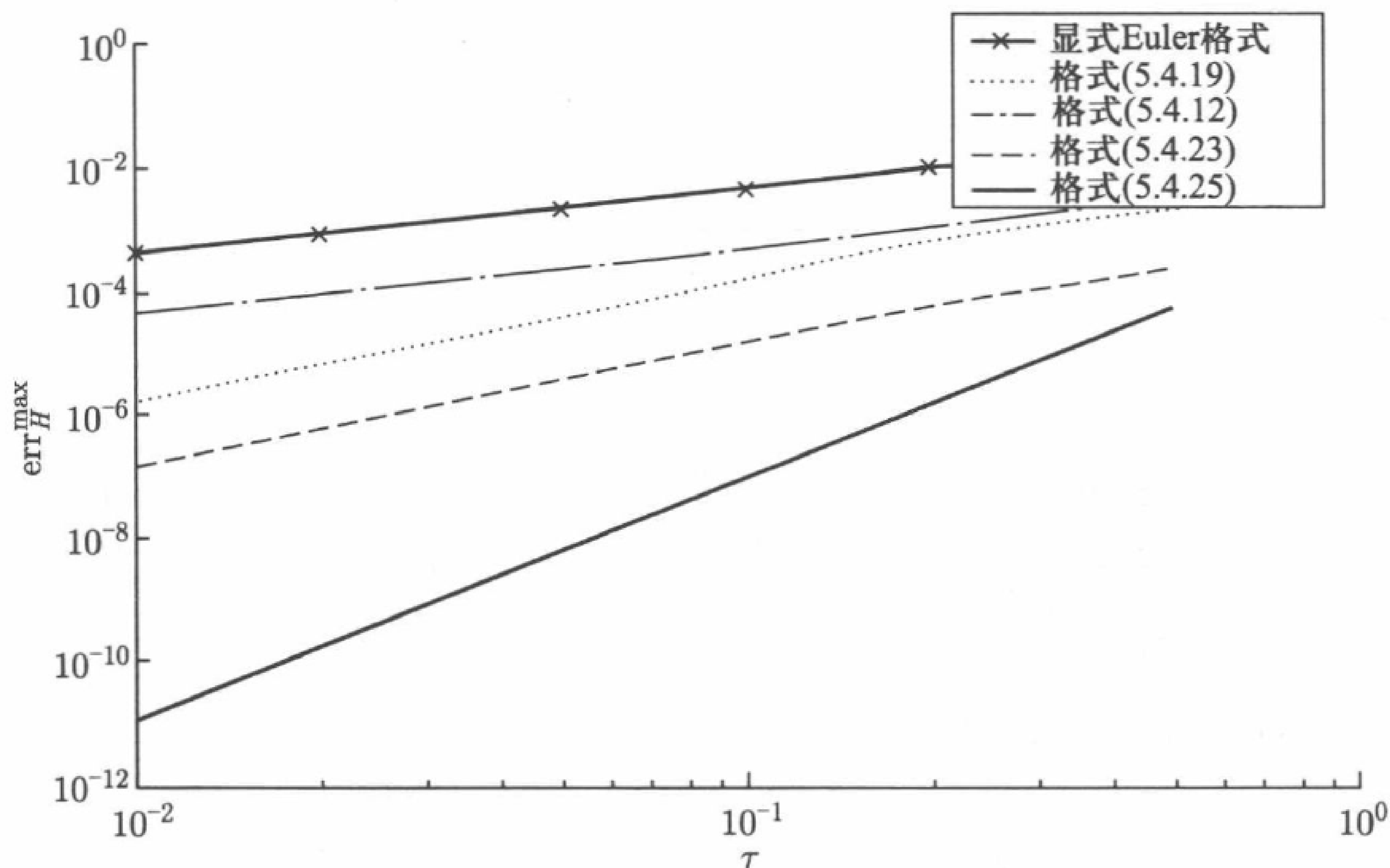


图 5.4.8 $\nu = 2.8$ 时不同差分格式的能量误差比较

附录 格式推导

考虑下面的方程:

$$\ddot{q} + \beta'(t)\dot{q} + V(q,t) = 0, \quad \text{即} \quad \begin{cases} \dot{p} + \beta'(t)p + V(q,t) = 0, \\ \dot{q} - p = 0, \end{cases}$$

或者更一般的形式

$$\begin{aligned} &\dot{p} + \beta'(t)p + V(q,t) = 0, \\ &\dot{q} - G(p,t) = 0. \end{aligned}$$

它有 Birkhoff 形式

$$\begin{pmatrix} 0 & -\mathrm{e}^{\beta(t)} \\ \mathrm{e}^{\beta(t)} & 0 \end{pmatrix} \begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} + \begin{pmatrix} -\beta'(t)\mathrm{e}^{\beta(t)}p - \mathrm{e}^{\beta(t)}V(q,t) \\ -\mathrm{e}^{\beta(t)}G(p,t) \end{pmatrix} = \mathbf{0}.$$

其 Birkhoff 函数 $\boldsymbol{F}$ 和 B 分别是

$$
\begin{aligned}
\boldsymbol{F} &= \left(\frac{1}{2}\mathrm{e}^{\beta(t)}p, -\frac{1}{2}\mathrm{e}^{\beta(t)}q\right)^{\mathrm{T}},\\
B &= \int_0^1 [(q,p)\cdot\nabla B(\lambda q,\lambda p)]\mathrm{d}\lambda,\\
\nabla B &= \begin{pmatrix} \dfrac{\partial B}{\partial q} \\ \dfrac{\partial B}{\partial p} \end{pmatrix} = -\boldsymbol{D} - \frac{\partial \boldsymbol{F}}{\partial t} = \begin{pmatrix} \dfrac{\beta'(t)}{2}\mathrm{e}^{\beta(t)}p + \mathrm{e}^{\beta(t)}V(q,t) \\ \dfrac{\beta'(t)}{2}\mathrm{e}^{\beta(t)}q + \mathrm{e}^{\beta(t)}G(p,t) \end{pmatrix},
\end{aligned}
$$

考虑 Darboux 变换 (Ⅲ 型)

$$
\boldsymbol{\alpha} = \begin{pmatrix} 0 & \mathrm{e}^{\frac{\beta(t)}{2}} & 0 & -\mathrm{e}^{\frac{\beta(t_0)}{2}} \\ -\mathrm{e}^{\frac{\beta(t)}{2}} & 0 & \mathrm{e}^{\frac{\beta(t_0)}{2}} & 0 \\ \frac{1}{2}\mathrm{e}^{\frac{\beta(t)}{2}} & 0 & \frac{1}{2}\mathrm{e}^{\frac{\beta(t_0)}{2}} & 0 \\ 0 & \frac{1}{2}\mathrm{e}^{\frac{\beta(t)}{2}} & 0 & \frac{1}{2}\mathrm{e}^{\frac{\beta(t_0)}{2}} \end{pmatrix}. \tag{5.5.1}
$$

令 $\begin{pmatrix} \widehat{\boldsymbol{w}} \\ \boldsymbol{w} \end{pmatrix} = \boldsymbol{\alpha}\begin{pmatrix} \widehat{\boldsymbol{z}} \\ \boldsymbol{z} \end{pmatrix}$, 其中

$$
\widehat{\boldsymbol{w}} = \begin{pmatrix} \widehat{Q} \\ \widehat{P} \end{pmatrix}, \quad \boldsymbol{w} = \begin{pmatrix} Q \\ P \end{pmatrix}, \quad \widehat{\boldsymbol{z}} = \begin{pmatrix} \widehat{q} \\ \widehat{p} \end{pmatrix}, \quad \boldsymbol{z} = \begin{pmatrix} q \\ p \end{pmatrix},
$$

得到

$$
\begin{aligned}
\widehat{Q} &= \mathrm{e}^{\frac{\beta(t)}{2}}\widehat{p} - \mathrm{e}^{\frac{\beta(t_0)}{2}}p,\\
\widehat{P} &= -\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{q} + \mathrm{e}^{\frac{\beta(t_0)}{2}}q,\\
Q &= \frac{1}{2}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{q} + \frac{1}{2}\mathrm{e}^{\frac{\beta(t_0)}{2}}q,\\
P &= \frac{1}{2}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{p} + \frac{1}{2}\mathrm{e}^{\frac{\beta(t_0)}{2}}p,
\end{aligned}
$$

其逆变换为

$$\begin{aligned}\widehat{q}&=\mathrm{e}^{-\frac{\beta(t)}{2}}Q-\frac{1}{2}\mathrm{e}^{-\frac{\beta(t)}{2}}\widehat{P},\\ \widehat{p}&=\mathrm{e}^{-\frac{\beta(t)}{2}}P+\frac{1}{2}\mathrm{e}^{-\frac{\beta(t)}{2}}\widehat{Q},\\ q&=\mathrm{e}^{-\frac{\beta(t_0)}{2}}Q+\frac{1}{2}\mathrm{e}^{-\frac{\beta(t_0)}{2}}\widehat{P},\\ p&=\mathrm{e}^{-\frac{\beta(t_0)}{2}}P-\frac{1}{2}\mathrm{e}^{-\frac{\beta(t_0)}{2}}\widehat{Q}.\end{aligned}$$

该变换对应的生成函数为

$$\phi_{\boldsymbol{w}}=\phi_{\boldsymbol{w}}^{(0)}+\phi_{\boldsymbol{w}}^{(1)}(t-t_0)+\phi_{\boldsymbol{w}}^{(2)}(t-t_0)^2+\cdots=\widehat{\boldsymbol{w}},$$

且

$$\frac{\partial\phi_{\boldsymbol{w}}}{\partial t}=\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}-\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}=\mathcal{A}(\widehat{\boldsymbol{w}},\boldsymbol{w},t).$$

将 $\mathcal{A}$ 在 $\left(\widehat{\boldsymbol{w}}\big|_{t=t_0}=\boldsymbol{w},t=t_0\right)$ 处展开, 首先得到

$$\phi_{\boldsymbol{w}}^{(0)}=\widehat{\boldsymbol{w}}\big|_{t=t_0}=\begin{pmatrix}0\\0\end{pmatrix},$$

进而有

$$\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\bigg|_{t=t_0}=\begin{pmatrix}0&0\\0&0\end{pmatrix},$$

$$\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}=\begin{pmatrix}\frac{\beta'(t)}{2}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{p}+\mathrm{e}^{\frac{\beta(t)}{2}}\dot{\widehat{p}}\\-\frac{\beta'(t)}{2}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{q}-\mathrm{e}^{\frac{\beta(t)}{2}}\dot{\widehat{q}}\end{pmatrix}=\begin{pmatrix}-\frac{\beta'(t)}{2}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{p}-\mathrm{e}^{\frac{\beta(t)}{2}}V(\widehat{q},t)\\-\frac{\beta'(t)}{2}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{q}-\mathrm{e}^{\frac{\beta(t)}{2}}G(\widehat{p},t)\end{pmatrix}$$

$$=\begin{pmatrix}-\frac{\beta'(t)}{2}P-\frac{\beta'(t)}{4}\widehat{Q}-\mathrm{e}^{\frac{\beta(t)}{2}}V\left(\mathrm{e}^{\frac{-\beta(t)}{2}}Q-\frac{1}{2}\mathrm{e}^{\frac{-\beta(t)}{2}}\widehat{P},t\right)\\-\frac{\beta'(t)}{2}Q+\frac{\beta'(t)}{4}\widehat{P}-\mathrm{e}^{\frac{\beta(t)}{2}}G\left(\mathrm{e}^{\frac{-\beta(t)}{2}}P+\frac{1}{2}\mathrm{e}^{\frac{-\beta(t)}{2}}\widehat{Q},t\right)\end{pmatrix},$$

$$\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\bigg|_{t=t_0}=\begin{pmatrix}-\frac{\beta'(t_0)}{2}P-\mathrm{e}^{\frac{\beta(t_0)}{2}}V\left(\mathrm{e}^{\frac{-\beta(t_0)}{2}}Q,t_0\right)\\-\frac{\beta'(t_0)}{2}Q-\mathrm{e}^{\frac{\beta(t_0)}{2}}G\left(\mathrm{e}^{\frac{-\beta(t_0)}{2}}P,t_0\right)\end{pmatrix},\tag{5.5.2}$$

$$\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}=\begin{pmatrix}\dfrac{\beta'(t)}{4}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{q}+\dfrac{1}{2}\mathrm{e}^{\frac{\beta(t)}{2}}\dot{\widehat{q}}\\\dfrac{\beta'(t)}{4}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{p}+\dfrac{1}{2}\mathrm{e}^{\frac{\beta(t)}{2}}\dot{\widehat{p}}\end{pmatrix}$$

$$=\begin{pmatrix}\dfrac{\beta'(t)}{4}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{q}+\dfrac{1}{2}\mathrm{e}^{\frac{\beta(t)}{2}}G(\widehat{p},t)\\-\dfrac{\beta'(t)}{4}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{p}-\dfrac{1}{2}\mathrm{e}^{\frac{\beta(t)}{2}}V(\widehat{q},t)\end{pmatrix}$$

$$=\begin{pmatrix}\dfrac{\beta'(t)}{4}Q-\dfrac{\beta'(t)}{8}\widehat{P}+\dfrac{1}{2}\mathrm{e}^{\frac{\beta(t)}{2}}G\left(\mathrm{e}^{\frac{-\beta(t)}{2}}P+\dfrac{1}{2}\mathrm{e}^{\frac{-\beta(t)}{2}}\widehat{Q},t\right)\\-\dfrac{\beta'(t)}{4}P-\dfrac{\beta'(t)}{8}\widehat{Q}-\dfrac{1}{2}\mathrm{e}^{\frac{\beta(t)}{2}}V\left(\mathrm{e}^{\frac{-\beta(t)}{2}}Q-\dfrac{1}{2}\mathrm{e}^{\frac{-\beta(t)}{2}}\widehat{P},t\right)\end{pmatrix},$$

$$\left.\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right|_{t=t_0}=\begin{pmatrix}\dfrac{\beta'(t_0)}{4}Q+\dfrac{1}{2}\mathrm{e}^{\frac{\beta(t_0)}{2}}G\left(\mathrm{e}^{\frac{-\beta(t_0)}{2}}P,t_0\right)\\-\dfrac{\beta'(t_0)}{4}P-\dfrac{1}{2}\mathrm{e}^{\frac{\beta(t_0)}{2}}V\left(\mathrm{e}^{\frac{-\beta(t_0)}{2}}Q,t_0\right)\end{pmatrix},\tag{5.5.3}$$

$$\phi_{\boldsymbol{w}}^{(1)}=\left.\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}-\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right)\right|_{t=t_0}\xlongequal{\left.\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\right|_{t=t_0}=\begin{pmatrix}0&0\\0&0\end{pmatrix}}\left.\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right|_{t=t_0}$$
$$=\text{方程 (5.5.2) 右端}.$$

因为
$$\left.\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\right|_{t=t_0}=\begin{pmatrix}0&0\\0&0\end{pmatrix},$$
所以
$$\left.\frac{\partial}{\partial t}\left(\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\right)\right|_{t=t_0}=\phi_{\boldsymbol{ww}}^{(1)}$$
$$=\begin{pmatrix}-V_q\left(\mathrm{e}^{\frac{-\beta(t_0)}{2}}Q,t_0\right)&-\dfrac{\beta'(t_0)}{2}\\-\dfrac{\beta'(t_0)}{2}&-G_p\left(\mathrm{e}^{\frac{-\beta(t_0)}{2}}P,t_0\right)\end{pmatrix}.\tag{5.5.4}$$

又
$$\phi_{\boldsymbol{w}}^{(2)}=\frac{1}{2}\left[\frac{\partial}{\partial t}\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right)+\frac{\partial}{\partial\widehat{\boldsymbol{w}}}\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right)\frac{\partial\widehat{\boldsymbol{w}}}{\partial t}-\frac{\partial}{\partial t}\left(\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\right)\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right.$$
$$\left.\left.-\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\left(\frac{\partial}{\partial t}\left(\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right)\right)-\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\frac{\partial}{\partial\widehat{\boldsymbol{w}}}\left(\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right)\frac{\partial\widehat{\boldsymbol{w}}}{\partial t}\right]\right|_{t=t_0},\tag{5.5.5}$$

则将 (5.5.4) 式代入 (5.5.5) 式, 得到

$$\phi_{\boldsymbol{w}}^{(2)} = \frac{1}{2}\left[\frac{\partial}{\partial t}\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right) + \frac{\partial}{\partial \widehat{\boldsymbol{w}}}\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right)\frac{\partial \widehat{\boldsymbol{w}}}{\partial t} - \frac{\partial}{\partial t}\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}\boldsymbol{w}}\right)\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right]\Bigg|_{t=t_0}$$

$$= \frac{1}{2}\left[\begin{pmatrix} -\frac{\beta''(t)}{2}P - \frac{\beta''(t)}{4}\widehat{Q} - \frac{\beta'(t)}{2}\mathrm{e}^{\frac{\beta(t)}{2}}V\left(\mathrm{e}^{\frac{-\beta(t)}{2}}Q - \frac{1}{2}\mathrm{e}^{\frac{-\beta(t)}{2}}\widehat{P}, t\right) \\ -\frac{\beta''(t)}{2}Q + \frac{\beta''(t)}{4}\widehat{P} - \frac{\beta'(t)}{2}\mathrm{e}^{\frac{\beta(t)}{2}}G\left(\mathrm{e}^{\frac{-\beta(t)}{2}}P + \frac{1}{2}\mathrm{e}^{\frac{-\beta(t)}{2}}\widehat{Q}, t\right) \end{pmatrix}\right.$$

$$+\frac{\beta'(t)}{2}\left(Q - \frac{1}{2}\widehat{P}\right)V_q\left(\mathrm{e}^{\frac{-\beta(t)}{2}}Q - \frac{1}{2}\mathrm{e}^{\frac{-\beta(t)}{2}}\widehat{P}, t\right)$$

$$+\frac{\beta'(t)}{2}\left(P + \frac{1}{2}\widehat{Q}\right)G_p\left(\mathrm{e}^{\frac{-\beta(t)}{2}}P + \frac{1}{2}\mathrm{e}^{\frac{-\beta(t)}{2}}\widehat{Q}, t\right)$$

$$-\mathrm{e}^{\frac{\beta(t)}{2}}V_t\left(\mathrm{e}^{\frac{-\beta(t)}{2}}Q - \frac{1}{2}\mathrm{e}^{\frac{-\beta(t)}{2}}\widehat{P}, t\right)$$

$$-\mathrm{e}^{\beta(t)}G_t\left(\mathrm{e}^{\frac{-\beta(t)}{2}}P + \frac{1}{2}\mathrm{e}^{\frac{-\beta(t)}{2}}\widehat{P}, t\right)\Bigg)$$

$$+\begin{pmatrix} -\frac{\beta'(t)}{4} & \frac{1}{2}V_q\left(\mathrm{e}^{\frac{-\beta(t)}{2}}Q - \frac{1}{2}\mathrm{e}^{\frac{-\beta(t)}{2}}\widehat{P}, t\right) \\ -\frac{1}{2}G_p\left(\mathrm{e}^{\frac{-\beta(t)}{2}}P + \frac{1}{2}\mathrm{e}^{\frac{-\beta(t)}{2}}\widehat{Q}, t\right) & \frac{\beta'(t)}{4} \end{pmatrix}$$

$$\cdot\begin{pmatrix} -\frac{\beta'(t_0)}{2}P - \mathrm{e}^{\frac{\beta(t_0)}{2}}V\left(\mathrm{e}^{\frac{-\beta(t_0)}{2}}Q, t_0\right) \\ -\frac{\beta'(t_0)}{2}Q - \mathrm{e}^{\frac{\beta(t_0)}{2}}G\left(\mathrm{e}^{\frac{-\beta(t_0)}{2}}P, t_0\right) \end{pmatrix}$$

$$-\begin{pmatrix} -V_q\left(\mathrm{e}^{\frac{-\beta(t_0)}{2}}Q, t_0\right) & -\frac{\beta'(t_0)}{2} \\ -\frac{\beta'(t_0)}{2} & -G_p\left(\mathrm{e}^{\frac{-\beta(t_0)}{2}}P, t_0\right) \end{pmatrix}$$

$$\left.\cdot\begin{pmatrix} \frac{\beta'(t_0)}{4}Q + \frac{1}{2}\mathrm{e}^{\frac{\beta(t_0)}{2}}G\left(\mathrm{e}^{\frac{-\beta(t_0)}{2}}P, t_0\right) \\ -\frac{\beta'(t_0)}{4}P - \frac{1}{2}\mathrm{e}^{\frac{\beta(t_0)}{2}}V\left(\mathrm{e}^{\frac{-\beta(t_0)}{2}}Q, t_0\right) \end{pmatrix}\right]\Bigg|_{t=t_0}.$$

整理得到

$$\phi_{\boldsymbol{w}}^{(2)}=\frac{1}{2}\left[\frac{\partial}{\partial t}\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right)\right.$$
$$+\begin{pmatrix}\frac{(\beta'(t_0))^2}{8}P+\frac{\beta'(t_0)}{4}\mathrm{e}^{\frac{\beta(t_0)}{2}}V-\frac{\beta'(t_0)}{4}QV_q-\frac{1}{2}\mathrm{e}^{\frac{\beta(t_0)}{2}}V_qG\\-\frac{(\beta'(t_0))^2}{8}Q-\frac{\beta'(t_0)}{4}\mathrm{e}^{\frac{\beta(t_0)}{2}}G+\frac{\beta'(t_0)}{4}PG_p+\frac{1}{2}\mathrm{e}^{\frac{\beta(t_0)}{2}}G_pV\end{pmatrix}$$
$$\left.\left.-\begin{pmatrix}-\frac{\beta'(t_0)}{4}QV_q-\frac{1}{2}\mathrm{e}^{\frac{\beta(t_0)}{2}}V_qG+\frac{(\beta'(t_0))^2}{8}P+\frac{\beta'(t_0)}{4}\mathrm{e}^{\frac{\beta(t_0)}{2}}V\\\frac{\beta'(t_0)}{4}PG_p+\frac{1}{2}\mathrm{e}^{\frac{\beta(t_0)}{2}}VG_p-\frac{(\beta'(t_0))^2}{8}Q-\frac{\beta'(t_0)}{4}\mathrm{e}^{\frac{\beta(t_0)}{2}}G\end{pmatrix}\right]\right|_{t=t_0}.$$

显然, 有

$$\begin{aligned}\phi_{\boldsymbol{w}}^{(2)}&=\frac{1}{2}\left.\left[\frac{\partial}{\partial t}\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right)\right]\right|_{t=t_0}\\&=\frac{1}{2}\begin{pmatrix}-\frac{\beta''(t_0)}{2}P-\frac{\beta'(t_0)}{2}\mathrm{e}^{\frac{\beta(t_0)}{2}}V\left(\mathrm{e}^{-\frac{\beta'(t_0)}{2}}Q,t_0\right)\\-\frac{\beta''(t_0)}{2}Q-\frac{\beta'(t_0)}{2}\mathrm{e}^{\frac{\beta(t_0)}{2}}G\left(\mathrm{e}^{-\frac{\beta'(t_0)}{2}}P,t_0\right)\\+\frac{\beta'(t_0)}{2}QV_q\left(\mathrm{e}^{-\frac{\beta(t_0)}{2}}Q,t_0\right)-\mathrm{e}^{\frac{\beta(t_0)}{2}}V_t\left(\mathrm{e}^{-\frac{\beta(t_0)}{2}}Q,t_0\right)\\+\frac{\beta'(t_0)}{2}PG_p\left(\mathrm{e}^{-\frac{\beta(t_0)}{2}}P,t_0\right)-\mathrm{e}^{\frac{\beta(t_0)}{2}}G_t\left(\mathrm{e}^{-\frac{\beta(t_0)}{2}}P,t_0\right)\end{pmatrix}.\end{aligned}\tag{5.5.6}$$

当 $m=1$ 次截断时, 有 1 阶精度的格式

$$\begin{aligned}\widehat{\boldsymbol{w}}=\phi_{\boldsymbol{w}}^{(1)}\tau&=\begin{pmatrix}\mathrm{e}^{\frac{\beta(t_{k+1})}{2}}p_{k+1}-\mathrm{e}^{\frac{\beta(t_k)}{2}}p_k\\-\mathrm{e}^{\frac{\beta(t_{k+1})}{2}}q_{k+1}+\mathrm{e}^{\frac{\beta(t_k)}{2}}q_k\end{pmatrix}\\&=\tau\begin{pmatrix}-\frac{\beta'(t_k)}{4}\mathrm{e}^{\frac{\beta(t_{k+1})}{2}}p_{k+1}-\frac{\beta'(t_k)}{4}\mathrm{e}^{\frac{\beta(t_k)}{2}}p_k\\-\frac{\beta'(t_k)}{4}\mathrm{e}^{\frac{\beta(t_{k+1})}{2}}q_{k+1}-\frac{\beta'(t_k)}{4}\mathrm{e}^{\frac{\beta(t_k)}{2}}q_k\\-\mathrm{e}^{\frac{\beta(t_k)}{2}}V\left(\frac{1}{2}\mathrm{e}^{\frac{\beta(t_{k+1})-\beta(t_k)}{2}}q_{k+1}+\frac{1}{2}q_k,t_k\right)\\-\mathrm{e}^{\frac{\beta(t_k)}{2}}G\left(\frac{1}{2}\mathrm{e}^{\frac{\beta(t_{k+1})-\beta(t_k)}{2}}p_{k+1}+\frac{1}{2}p_k,t_k\right)\end{pmatrix}.\end{aligned}\tag{5.5.7}$$

当 $m=2$ 次截断时, 有 2 阶精度的格式

$$\widehat{\boldsymbol{w}}=\phi_{\boldsymbol{w}}^{(1)}\tau+\phi_{\boldsymbol{w}}^{(2)}\tau^2,$$

即在方程 (5.5.7) 的右边加上 $\phi_{\boldsymbol{w}}^{(2)}\tau^2$, 其中

$$\phi_{\boldsymbol{w}}^{(2)}=\begin{pmatrix} -\dfrac{\beta''(t_k)}{4}\mathrm{e}^{\frac{\beta(t_{k+1})}{2}}p_{k+1}-\dfrac{\beta''(t_k)}{4}\mathrm{e}^{\frac{\beta(t_k)}{2}}p_k-\dfrac{\beta'(t_k)}{2}\mathrm{e}^{\frac{\beta(t_k)}{2}}V \\ -\dfrac{\beta''(t_k)}{4}\mathrm{e}^{\frac{\beta(t_{k+1})}{2}}q_{k+1}-\dfrac{\beta''(t_k)}{4}\mathrm{e}^{\frac{\beta(t_k)}{2}}q_k-\dfrac{\beta'(t_k)}{2}\mathrm{e}^{\frac{\beta(t_k)}{2}}G \\ +\left(\dfrac{\beta'(t_k)}{4}\mathrm{e}^{\frac{\beta(t_{k+1})}{2}}q_{k+1}+\dfrac{\beta'(t_k)}{4}\mathrm{e}^{\frac{\beta(t_k)}{2}}q_k\right)V_q-\mathrm{e}^{\frac{\beta(t_k)}{2}}V_t \\ +\left(\dfrac{\beta'(t_k)}{4}\mathrm{e}^{\frac{\beta(t_{k+1})}{2}}p_{k+1}+\dfrac{\beta'(t_k)}{4}\mathrm{e}^{\frac{\beta(t_k)}{2}}p_k\right)G_p-\mathrm{e}^{\frac{\beta(t_k)}{2}}G_t \end{pmatrix},$$

这里

$$V_\gamma=V_\gamma\left(\frac{1}{2}\mathrm{e}^{\frac{\beta(t_{k+1})-\beta(t_k)}{2}}q_{k+1}+\frac{1}{2}q_k,t_k\right),\quad V_\gamma=\begin{cases}V,\\V_q,\\V_t,\end{cases}$$

$$G_\gamma=G_\gamma\left(\frac{1}{2}\mathrm{e}^{\frac{\beta(t_{k+1})-\beta(t_k)}{2}}p_{k+1}+\frac{1}{2}p_k,t_k\right),\quad G_\gamma=\begin{cases}G,\\G_p,\\G_t.\end{cases}$$

当 $V(q,t)=q, G(p,t)=p$ 时, $V_q=1, V_t=0, G_p=1, G_t=0$, 从而

$$\phi_{\boldsymbol{w}}^{(1)}=\begin{pmatrix}-\dfrac{\beta'(t_0)}{2}P-Q\\ -\dfrac{\beta'(t_0)}{2}Q-P\end{pmatrix},$$

$$\phi_{\boldsymbol{w}}^{(2)}=\begin{pmatrix}-\dfrac{\beta''(t_0)}{2}P-\dfrac{\beta'(t_0)}{2}Q+\dfrac{\beta'(t_0)}{2}Q\\ -\dfrac{\beta''(t_0)}{2}Q-\dfrac{\beta'(t_0)}{2}P+\dfrac{\beta'(t_0)}{2}P\end{pmatrix}=\begin{pmatrix}\dfrac{-\beta''(t_0)}{2}P\\ \dfrac{-\beta''(t_0)}{2}Q\end{pmatrix}.$$

这和直接由线性方程 $\dot{q}+\beta(t)\dot{q}+q=0$ 出发得到的 $\phi_{\boldsymbol{w}}^{(1)},\phi_{\boldsymbol{w}}^{(2)}$ 完全一致.

下面由线性方程出发推导格式, 过程如下:

(1) 针对方程 $\dot{q}+\beta(t)\dot{q}+q=0$ 的格式推导.

该线性方程等价于

$$\begin{cases}\dot{p}+\beta'(+)p+q=0,\\ \dot{q}-p=0.\end{cases}$$

依然考虑 Darboux 变换 (5.5.1). 显然

$$\widehat{\boldsymbol{w}}\big|_{t=t_0}=\begin{pmatrix}0\\0\end{pmatrix},$$

而生成函数为

$$\phi_{\boldsymbol{w}}=\phi_{\boldsymbol{w}}^{(0)}+\phi_{\boldsymbol{w}}^{(1)}(t-t_0)+\phi_{\boldsymbol{w}}^{(2)}(t-t_0)^2+\cdots=\widehat{\boldsymbol{w}},$$

所以

$$\phi_{\boldsymbol{w}}^{(0)}=\widehat{\boldsymbol{w}}\big|_{t=t_0}=\begin{pmatrix}0\\0\end{pmatrix},$$

$$\frac{\partial\phi_{\boldsymbol{w}}}{\partial t}=\phi_{\boldsymbol{w}}^{(1)}+2\phi_{\boldsymbol{w}}^{(2)}(t-t_0)+3\phi_{\boldsymbol{w}}^{(3)}(t-t_0)^2+\cdots,$$

又因为

$$\frac{\partial\phi_{\boldsymbol{w}}}{\partial t}=\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}-\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}=\mathcal{A}(\widehat{\boldsymbol{w}},\boldsymbol{w},t),$$

所以将 $\mathcal{A}(\widehat{\boldsymbol{w}},\boldsymbol{w},t)$ 在 $\widehat{\boldsymbol{w}}\big|_{t=t_0}$ 和 $t=t_0$ 处展开, 与 $\dfrac{\partial\phi_{\boldsymbol{w}}}{\partial t}$ 比较, 就能求出 $\phi_{\boldsymbol{w}}^{(i)}$. 我们有

$$\phi_{\boldsymbol{w}}^{(1)}=\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\bigg|_{t=t_0}-\left(\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right)\bigg|_{t=t_0}.\tag{5.5.8}$$

首先由变换 (5.5.1) 找到 $\dfrac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t},\dfrac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}$:

$$\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}=\begin{pmatrix}\dfrac{\beta'(t)}{2}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{p}+\mathrm{e}^{\frac{\beta(t)}{2}}\dot{\widehat{p}}\\ -\dfrac{\beta'(t)}{2}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{q}-\mathrm{e}^{\frac{\beta(t)}{2}}\dot{\widehat{q}}\end{pmatrix}=\begin{pmatrix}-\dfrac{\beta'(t)}{2}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{p}-\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{q}\\ -\dfrac{\beta'(t)}{2}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{q}-\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{p}\end{pmatrix}$$

$$=\begin{pmatrix}-\dfrac{\beta'(t)}{2}P-\dfrac{\beta'(t)}{4}\widehat{Q}-Q+\dfrac{1}{2}\widehat{P}\\ -\dfrac{\beta'(t)}{2}Q+\dfrac{\beta'(t)}{4}\widehat{P}-P-\dfrac{1}{2}\widehat{Q}\end{pmatrix},$$

$$\left.\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right|_{t=t_0}=\begin{pmatrix}-\dfrac{\beta'(t_0)}{2}P-Q\\-\dfrac{\beta'(t_0)}{2}Q-P\end{pmatrix},$$

$$\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}=\begin{pmatrix}\dfrac{\beta'(t)}{4}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{q}+\dfrac{1}{2}\mathrm{e}^{\frac{\beta(t)}{2}}\dot{\widehat{q}}\\\dfrac{\beta'(t)}{4}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{p}+\dfrac{1}{2}\mathrm{e}^{\frac{\beta(t)}{2}}\dot{\widehat{p}}\end{pmatrix}=\begin{pmatrix}\dfrac{\beta'(t)}{4}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{q}+\dfrac{1}{2}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{p}\\-\dfrac{\beta'(t)}{4}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{p}-\dfrac{1}{2}\mathrm{e}^{\frac{\beta(t)}{2}}\widehat{q}\end{pmatrix}$$

$$=\begin{pmatrix}\dfrac{\beta'(t)}{4}Q-\dfrac{\beta'(t)}{8}\widehat{P}+\dfrac{1}{2}P+\dfrac{1}{4}\widehat{Q}\\-\dfrac{\beta'(t)}{4}P-\dfrac{\beta'(t)}{8}\widehat{Q}-\dfrac{1}{2}Q+\dfrac{1}{4}\widehat{P}\end{pmatrix},$$

$$\left.\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right|_{t=t_0}=\begin{pmatrix}\dfrac{\beta'(t_0)}{4}Q+\dfrac{1}{2}P\\-\dfrac{\beta'(t_0)}{2}P-\dfrac{1}{2}Q\end{pmatrix}.$$

将以上格式代入 (5.5.8) 式, 得到

$$\phi_{\boldsymbol{w}}^{(1)}=\left.\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right|_{t=t_0}=\begin{pmatrix}-\dfrac{\beta'(t_0)}{2}P-Q\\-\dfrac{\beta'(t_0)}{2}Q-P\end{pmatrix},$$

进而有

$$\phi_{\boldsymbol{w}}^{(2)}=\frac{1}{2}\frac{\partial}{\partial t}\left.\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}-\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right)\right|_{t=t_0}$$

$$=\frac{1}{2}\left[\frac{\partial}{\partial t}\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right)+\frac{\partial}{\partial\widehat{\boldsymbol{w}}}\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right)\frac{\partial\widehat{\boldsymbol{w}}}{\partial t}-\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\frac{\partial}{\partial\widehat{\boldsymbol{w}}}\left(\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right)\frac{\partial\widehat{\boldsymbol{w}}}{\partial t}\right.$$

$$\left.\left.-\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\frac{\partial}{\partial t}\left(\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right)-\frac{\partial}{\partial t}\left(\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\right)\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right]\right|_{t=t_0}.$$

因为

$$\left.\frac{\partial\widehat{\boldsymbol{w}}}{\partial t}\right|_{t=t_0}=\phi_{\boldsymbol{w}}^{(1)},$$

$$\left.\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\right|_{t=t_0}=\phi^{(0)}_{\boldsymbol{ww}}=\begin{pmatrix}0&0\\0&0\end{pmatrix},$$

$$\left.\frac{\partial}{\partial t}\left(\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{\omega}}\right)\right|_{t=t_0}=\phi^{(1)}_{\boldsymbol{ww}}=\begin{pmatrix}-1&-\dfrac{\beta'(t_0)}{2}\\-\dfrac{\beta'(t_0)}{2}&-1\end{pmatrix},$$

所以

$$\phi^{(2)}_{\boldsymbol{w}}=\frac{1}{2}\left.\left[\frac{\partial}{\partial t}\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right)+\frac{\partial}{\partial\widehat{\boldsymbol{w}}}\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right)\frac{\partial\widehat{\boldsymbol{w}}}{\partial t}-\frac{\partial}{\partial t}\left(\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\right)\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t}\right]\right|_{t=t_0}$$

$$=\frac{1}{2}\left[\begin{pmatrix}-\dfrac{\beta''(t)}{2}P-\dfrac{\beta''(t)}{2}\widehat{Q}\\-\dfrac{\beta''(t)}{2}Q-\dfrac{\beta''(t)}{2}\widehat{P}\end{pmatrix}+\begin{pmatrix}-\dfrac{\beta'(t_0)}{4}&\dfrac{1}{2}\\-\dfrac{1}{2}&\dfrac{\beta'(t_0)}{4}\end{pmatrix}\right.$$

$$\cdot\begin{pmatrix}-\dfrac{\beta'(t_0)}{2}P-Q\\-\dfrac{\beta'(t_0)}{2}Q-P\end{pmatrix}-\begin{pmatrix}-1&-\dfrac{\beta'(t_0)}{2}\\\dfrac{-\beta'(t_0)}{2}&-1\end{pmatrix}$$

$$\left.\left.\cdot\begin{pmatrix}-\dfrac{\beta'(t_0)}{4}Q+\dfrac{1}{2}P\\\dfrac{-\beta'(t_0)}{4}P-\dfrac{1}{2}Q\end{pmatrix}\right]\right|_{t=t_0}$$

$$=\frac{1}{2}\left[\frac{\partial}{\partial t}\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right)+\begin{pmatrix}\dfrac{(\beta'(t_0))^2}{8}P+\dfrac{\beta'(t_0)}{4}Q-\dfrac{\beta'(t_0)}{4}Q-\dfrac{P}{2}\\\dfrac{-(\beta'(t_0))^2}{8}Q-\dfrac{\beta'(t_0)}{4}P+\dfrac{\beta'(t_0)}{4}P+\dfrac{Q}{2}\end{pmatrix}\right.$$

$$\left.\left.-\begin{pmatrix}\dfrac{(\beta'(t_0))^2}{8}P+\dfrac{\beta'(t_0)}{4}Q-\dfrac{\beta'(t_0)}{4}Q-\dfrac{P}{2}\\\dfrac{-(\beta'(t_0))^2}{8}Q-\dfrac{\beta'(t_0)}{4}P+\dfrac{\beta'(t_0)}{4}P+\dfrac{Q}{2}\end{pmatrix}\right]\right|_{t=t_0}$$

$$=\frac{1}{2}\left.\left[\frac{\partial}{\partial t}\left(\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right)\right]\right|_{t=t_0}=\begin{pmatrix}\dfrac{-\beta''(t_0)}{4}P\\\dfrac{-\beta''(t_0)}{4}Q\end{pmatrix}.$$

最终有

$$\phi_{\boldsymbol{w}}^{(1)} = \begin{pmatrix} \dfrac{-\beta'(t_0)}{2}P - Q \\ \dfrac{-\beta'(t_0)}{2}Q - P \end{pmatrix}, \quad \phi_{\boldsymbol{w}}^{(2)} = \begin{pmatrix} \dfrac{-\beta''(t_0)}{4}P \\ \dfrac{-\beta''(t_0)}{4}Q \end{pmatrix}.$$

(2) 对称格式 (5.4.23) 的推导.

这时有

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \mathrm{e}^{\frac{\nu t}{2}} & 0 & -\mathrm{e}^{\frac{\nu t_0}{2}} \\ -\mathrm{e}^{\frac{\nu t}{2}} & 0 & \mathrm{e}^{\frac{\nu t_0}{2}} & 0 \\ \dfrac{1}{2}\mathrm{e}^{\frac{\nu t}{2}} & 0 & \dfrac{1}{2}\mathrm{e}^{\frac{\nu t_0}{2}} & 0 \\ 0 & \dfrac{1}{2}\mathrm{e}^{\frac{\nu t}{2}} & 0 & \dfrac{1}{2}\mathrm{e}^{\frac{\nu t_0}{2}} \end{pmatrix}, \quad \begin{pmatrix} \widehat{Q} \\ \widehat{P} \\ Q \\ P \end{pmatrix} = \boldsymbol{\alpha} \begin{pmatrix} \widehat{q} \\ \widehat{p} \\ q \\ p \end{pmatrix},$$

$$\begin{pmatrix} \widehat{Q} \\ \widehat{P} \end{pmatrix} = \begin{pmatrix} \mathrm{e}^{\frac{\nu t}{2}}\widehat{p} - \mathrm{e}^{\frac{\nu t_0}{2}}p \\ -\mathrm{e}^{\frac{\nu t}{2}}\widehat{q} + \mathrm{e}^{\frac{\nu t_0}{2}}q \end{pmatrix}, \quad \begin{pmatrix} Q \\ P \end{pmatrix} = \begin{pmatrix} \dfrac{1}{2}\mathrm{e}^{\frac{\nu t}{2}}\widehat{q} + \dfrac{1}{2}\mathrm{e}^{\frac{\nu t_0}{2}}q \\ \dfrac{1}{2}\mathrm{e}^{\frac{\nu t}{2}}\widehat{p} + \dfrac{1}{2}\mathrm{e}^{\frac{\nu t_0}{2}}p \end{pmatrix},$$

$$\boldsymbol{\alpha}_*^{-1} = \begin{pmatrix} 0 & -\dfrac{1}{2}\mathrm{e}^{\frac{-\nu t}{2}} & \mathrm{e}^{\frac{-\nu t}{2}} & 0 \\ \dfrac{1}{2}\mathrm{e}^{\frac{-\nu t}{2}} & 0 & 0 & \mathrm{e}^{\frac{-\nu t}{2}} \\ 0 & \dfrac{1}{2}\mathrm{e}^{\frac{-\nu t_0}{2}} & \mathrm{e}^{\frac{-\nu t_0}{2}} & 0 \\ -\dfrac{1}{2}\mathrm{e}^{\frac{-\nu t_0}{2}} & 0 & 0 & \mathrm{e}^{\frac{-\nu t_0}{2}} \end{pmatrix},$$

$$\widehat{q} = -\frac{1}{2}\mathrm{e}^{-\frac{\nu t}{2}}\widehat{P} + \mathrm{e}^{-\frac{\nu t_0}{2}}Q,$$

$$\widehat{p} = \frac{1}{2}\mathrm{e}^{-\frac{\nu t}{2}}\widehat{Q} + \mathrm{e}^{-\frac{\nu t_0}{2}}P,$$

$$q = \frac{1}{2}\mathrm{e}^{-\frac{vt_0}{2}}\widehat{P} + \mathrm{e}^{-\frac{vt_0}{2}}Q,$$

$$p = -\frac{1}{2}\mathrm{e}^{-\frac{vt_0}{2}}\widehat{Q} + \mathrm{e}^{-\frac{vt_0}{2}}P,$$

$$\widehat{\boldsymbol{w}} = \left.\begin{pmatrix} \mathrm{e}^{\frac{\nu t}{2}}\widehat{p} - \mathrm{e}^{\frac{\nu t_0}{2}} p \\ -\mathrm{e}^{\frac{\nu t}{2}}\widehat{q} + \mathrm{e}^{\frac{\nu t_0}{2}} q \end{pmatrix}\right|_{t=t_0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$\begin{aligned}
\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t} &= \begin{pmatrix} \frac{\nu}{2}\mathrm{e}^{\frac{\nu t}{2}}\widehat{p} + \mathrm{e}^{\frac{\nu t}{2}}\dot{\widehat{p}} \\ -\frac{\nu}{2}\mathrm{e}^{\frac{\nu t}{2}}\widehat{q} - \mathrm{e}^{\frac{\nu t}{2}}\dot{\widehat{q}} \end{pmatrix} = \begin{pmatrix} \frac{\nu}{2}\mathrm{e}^{\frac{\nu t}{2}}\widehat{p} - \nu\mathrm{e}^{\frac{\nu t}{2}}\widehat{p} - \mathrm{e}^{\frac{\nu t}{2}}\widehat{q} \\ -\frac{\nu}{2}\mathrm{e}^{\frac{\nu t}{2}}\widehat{q} - \mathrm{e}^{\frac{\nu t}{2}}\widehat{p} \end{pmatrix} \\
&= \begin{pmatrix} -\frac{\nu}{2}\mathrm{e}^{\frac{\nu t}{2}}\widehat{p} - \mathrm{e}^{\frac{\nu t}{2}}\widehat{q} \\ -\frac{\nu}{2}\mathrm{e}^{\frac{\nu t}{2}}\widehat{q} - \mathrm{e}^{\frac{\nu t}{2}}\widehat{p} \end{pmatrix} = \begin{pmatrix} -\frac{\nu}{4}\widehat{Q} - \frac{\nu}{2}P + \frac{1}{2}\widehat{P} - Q \\ \frac{\nu}{4}\widehat{P} - \frac{\nu}{2}Q - \frac{1}{2}\widehat{Q} - P \end{pmatrix}, \\
\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}t} &= \begin{pmatrix} \frac{\nu}{4}\mathrm{e}^{\frac{\nu t}{2}}\widehat{q} + \frac{1}{2}\mathrm{e}^{\frac{\nu t}{2}}\dot{\widehat{q}} \\ \frac{\nu}{4}\mathrm{e}^{\frac{\nu t}{2}}\widehat{p} + \frac{1}{2}\mathrm{e}^{\frac{\nu t}{2}}\dot{\widehat{p}} \end{pmatrix} = \begin{pmatrix} \frac{\nu}{4}\mathrm{e}^{\frac{\nu t}{2}}\widehat{q} + \frac{1}{2}\mathrm{e}^{\frac{\nu t}{2}}\widehat{p} \\ \frac{\nu}{4}\mathrm{e}^{\frac{\nu t}{2}}\widehat{p} - \frac{\nu}{2}\mathrm{e}^{\frac{\nu t}{2}}\widehat{p} - \frac{1}{2}\mathrm{e}^{\frac{\nu t}{2}}\widehat{q} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\nu}{4}\mathrm{e}^{\frac{\nu t}{2}}\widehat{q} + \frac{1}{2}\mathrm{e}^{\frac{\nu t}{2}}\widehat{p} \\ -\frac{\nu}{4}\mathrm{e}^{\frac{\nu t}{2}}\widehat{p} - \frac{1}{2}\mathrm{e}^{\frac{\nu t}{2}}\widehat{q} \end{pmatrix} = \begin{pmatrix} -\frac{\nu}{8}\widehat{P} + \frac{\nu}{4}Q + \frac{1}{4}\widehat{Q} + \frac{1}{2}P \\ -\frac{\nu}{8}\widehat{Q} - \frac{\nu}{4}P + \frac{1}{4}\widehat{P} - \frac{1}{2}Q \end{pmatrix}.
\end{aligned} \tag{5.5.9}$$

因为

$$\left.\frac{\partial\widehat{\boldsymbol{w}}}{\partial\boldsymbol{w}}\right|_{t=t_0} = \phi_{\boldsymbol{ww}}^{(0)} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

所以

$$\phi_{\boldsymbol{w}}^{(1)} = \left.\frac{\mathrm{d}\widehat{\boldsymbol{w}}}{\mathrm{d}t}\right|_{t=t_0} = \begin{pmatrix} -\frac{\nu}{2}P - Q \\ -\frac{\nu}{2}Q - P \end{pmatrix}. \tag{5.5.10}$$

由 $\widehat{\boldsymbol{w}} = \phi_{\boldsymbol{w}}^{(1)}\tau$ 得到格式 (5.4.23).

参 考 文 献

[Arn89]　Arnold V I. *Mathematical Methods of Classical Mechanics.* 2nd edition. Berlin: Springer-Verlag, 1989.

[Fen86]　Feng K. Difference schemes for Hamiltonian formalism and symplectic geometry. *J Comput Math*, 1986, 4: 279–289.

[FQ87] Feng K, Qin M Z. The symplectic methods for the computation of Hamiltonian equations. *Numerical Methods for Partial Differential Equations: Lecture Notes in Mathematics*, Berlin: Springer, 1987, 1297: 1–37.

[FWQW89] Feng K, Wu H M, Qin M Z, Wang D L. Construction of canonical difference schemes for Hamiltonian formalism via generating functions. *J Comput Math*, 1989, 7: 71–96.

[Mos62] Moser J. On invariant curves of area-preserving mappings of an annulus. *Nachr Akad Wiss Gottingen II Math Phys*, 1962: 1–20.

[MW01] Marsden J E, West M. Discrete mechanics and variational integrators. *Acta Numerica*, 2001, 10: 357–514.

[QZ92] Qin M Z, Zhu W J. Construction of higher order symplectic schemes by composition. *Computing*, 1992, 47: 309–321.

[San83a] Santilli R M. *Foundations of Theoretical Mechanics I*. 2nd edition. New York: Springer-Verlag, 1983.

[San83b] Santilli R M. *Foundations of Theoretical Mechanics II*. 2nd edition. New York: Springer-Verlag, 1983.

[SQ04] Su H L, Qin M Z. Symplectic schemes for Birkhoffian system. *Commun Theor Phys*, 2004, 41(3): 329–334.

[Wan91] Wang D L. Poisson difference schemes for Hamiltonian systems on Poisson manifolds. *J Comput Math*, 1991, 9: 115–124.

[Yos90] Yoshida H. Construction of higher order symplectic integrators. *Phys Lett A*, 1990, 150: 262–268.

[梅凤翔13] 梅凤翔. 广义 Birkhoff 系统动力学. 北京: 科学出版社, 2013.

第六章 偏微分 Birkhoff 系统的多辛结构及多辛格式

在第五章中, 我们讨论了有限维 Birkhoff 辛结构下受外力作用的系统的几何结构和生成函数构造法. 有限维 Hamilton 系统是包含在半自治和自治的 Birkhoff 系统中的. 对 Hamilton 偏微分方程的多辛形式的描述, 引出了对保持多辛守恒律的几何数值方法的研究 (如果是自治的 Hamilton 系统, 其几何数值方法的研究也包括如何保持能量守恒律). 多辛 Hamilton 形式首先是由 Bridges 于 1997 年提出[Bri97a], 用以研究波动方程的稳定性, 后来又发展成一种完整的研究多辛算法的方法. 简单地讲, 多辛方法就是给出一个系统或者一个偏微分方程组, 首先判断它的对称性, 或者说看它是不是 Euler-Lagrange 方程, 假如是, 那么利用势函数公式求出它的 Lagrange 函数, 再利用共轭 Legendre 变换导出它的多辛 Hamilton 形式, 并且证明由中点离散多辛 Hamilton 形式[BR01,REi00] 可以得到多辛格式. 另外, Marsden 等人从 Lagrange 泛函出发, 利用变分求不动点, 从通常所谓的边界项定义出一个多辛结构, 并证明变分问题的 Lagrange 流是多辛守恒的, 同时在离散空间上进行了以上讨论, 从而给出了多辛守恒的变分积分子概念[MPS98]. 无论是 Marsden 还是 Bridges, 他们都是从二阶或二阶以上通常意义的对称系统即 Euler-Lagrange 方程出发讨论的, 因为对于单个方程, 如果最高阶导数是奇数阶的, 那么它们必然不对称, 所以一阶方程就不在他们的讨论范围内. 但是, 我们发现一阶方程组中, 除了多辛 Hamilton 方程之外, 还有一类带有耗散项的方程组也是对称的. 此类系统就是我们在本章将要学习的多辛 Birkhoff 系统. 本章我们从偏微分方程出发, 通过方程的对称性讨论其几何结构, 进而研究其保结构算法.

§6.1 多辛 Hamilton 方程及其推广

多辛 Hamilton 形式使得 Hamilton 方程在每个自变量方向都有一个以反对称矩阵为代表的辛结构, 我们有时也称之为多辛 Bridges 形式. Birkhoff 常微分方程首次由 Santilli 提出, 他在文献 [San83] 中讨论了 Birkhoff 常微分方程、它的代数和几何框架以及能量函数等. 我们将结合多辛 Hamilton 方程和 Birkhoff 常微分方程, 提出多辛 Birkhoff 偏微分方程的概念, 并研究它的对称性和多辛守恒律. 为了叙述简便, 我们只考虑 $1+1$ 维的偏微分方程, 对于高维系统, 可类似推广.

首先回顾一下多辛 Hamilton 方程. 所谓的多辛就是 Hamilton 系统在时间和空间方向分别存在一个辛结构. 辛结构的局部化, 也就是将偏微分方程的无限维辛 Hamilton 形式变为有限维多辛形式. 多辛 Hamilton 方程概念可以简述如下: 令 $\boldsymbol{M}$ 和 $\boldsymbol{K}$ 是 $\mathbf{R}^n\,(n\geqslant 3)$ 上的任意反对称矩阵, S 是 $\mathbf{R}^n\times\mathbf{R}^2$ 到 $\mathbf{R}$ 上的任意光滑函数. 一个方程若是能表示成下面的形式:

$$\boldsymbol{M}\boldsymbol{z}_t+\boldsymbol{K}\boldsymbol{z}_x=\nabla_{\boldsymbol{z}}S(\boldsymbol{z},x,t),\quad \boldsymbol{z}\in\mathbf{R}^n,\ (x,t)\in\mathbf{R}^2, \tag{6.1.1}$$

则称该方程是多辛 Hamilton 形式 (6.1.1) 下的**多辛 Hamilton 方程**, 这里的梯度算子 $\nabla_{\boldsymbol{z}}$ 利用 $\mathbf{R}^n$ 中的标准内积来定义. $\mathbf{R}^n$ 中的标准内积记为 $\langle\cdot,\cdot\rangle$. 矩阵 $\boldsymbol{M}$ 和 $\boldsymbol{K}$ 可以定义两个辛 2-形式, 分别是

$$\omega(U,V)=\langle\boldsymbol{M}U,V\rangle,\quad \kappa(U,V)=\langle\boldsymbol{K}U,V\rangle, \tag{6.1.2}$$

Bridges 称这两个辛 2-形式为 Pre-symplectic 形式. 这两个 2-形式又分别在 $\mathbf{R}^m\,(m=\mathrm{rank}(\boldsymbol{M})\leqslant n)$ 和 $\mathbf{R}^k\,(k=\mathrm{rank}(\boldsymbol{K})\leqslant n)$ 上定义了两个辛结构. 形式对 (ω,κ) 满足多辛守恒律

$$\frac{\mathrm{d}\omega}{\mathrm{d}t}+\frac{\mathrm{d}\kappa}{\mathrm{d}x}=0. \tag{6.1.3}$$

系统 (6.1.1) 的一个多辛格式就是对它自身的一个数值逼近, 例如

$$\boldsymbol{M}\partial t^{i,j}\boldsymbol{z}_{i,j}+\boldsymbol{K}\partial x^{i,j}\boldsymbol{z}_{i,j}=(\nabla_{\boldsymbol{z}}S(\boldsymbol{z}_{i,j},x_i,t_j))_{i,j},\quad \boldsymbol{z}_{i,j}=\boldsymbol{z}(x_i,t_j), \tag{6.1.4}$$

其中 $\partial t^{i,j}$ 和 $\partial x^{i,j}$ 分别表示 $\dfrac{\partial}{\partial t}$ 和 $\dfrac{\partial}{\partial x}$ 在 (x_i, t_j) 处的离散近似. 这个数值逼近保持离散的多辛守恒律:

$$\partial t^{i,j}\omega_{i,j} + \partial x^{i,j}\kappa_{i,j} = 0. \tag{6.1.5}$$

假如函数 S 不显含时间 t 和空间变量 x, 系统 (6.1.1) 就是自治的多辛 Hamilton 系统, 这样的系统除了有多辛守恒律之外, 还有形式能量和动量守恒律:

$$\partial_t\mathcal{E}(\boldsymbol{z}) + \partial_x\mathcal{F}(\boldsymbol{z}) = 0 \quad 和 \quad \partial_t\mathcal{I}(\boldsymbol{z}) + \partial_x\mathcal{G}(\boldsymbol{z}) = 0 \tag{6.1.6}$$

分别是系统 (6.1.1) 的形式能量和动量守恒律, 其中

$$\begin{aligned}\mathcal{E}(\boldsymbol{z}) &= S(\boldsymbol{z}) - \frac{1}{2}\kappa(\boldsymbol{z}_x, \boldsymbol{z}), \quad \mathcal{F}(\boldsymbol{z}) = \frac{1}{2}\kappa(\boldsymbol{z}_t, \boldsymbol{z}),\\ \mathcal{G}(\boldsymbol{z}) &= S(\boldsymbol{z}) - \frac{1}{2}\omega(\boldsymbol{z}_t, \boldsymbol{z}), \quad \mathcal{I}(\boldsymbol{z}) = \frac{1}{2}\omega(\boldsymbol{z}_x, \boldsymbol{z}).\end{aligned}$$

系统 (6.1.1) 的多辛格式同样保持离散的形式能量和动量守恒. 非自治的多辛 Hamilton 系统, 也就是势函数 S 显含自变量 (x, t) 时, 不存在明显的能量和动量守恒律. 不过即使作为能量耗散系统, 有些系统的能量耗散也有一定的规律, 所以我们也尝试找到它们的形式能量守恒律.

如果在构形空间上考虑, 把自变量和因变量等价地看待, 我们可以很自然地把多辛 Hamilton 系统推广到更一般的形式. 令 $\boldsymbol{M}(\boldsymbol{z},x,t)$ 和 $\boldsymbol{K}(\boldsymbol{z}, x, t)$ 也是 $\mathbf{R}^n \times \mathbf{R}^2$ $(n \geqslant 3)$ 上的反对称矩阵, B 是从 $\mathbf{R}^n \times \mathbf{R}^2$ 到 $\mathbf{R}$ 上的光滑函数, 那么一般的多辛 Hamilton 系统可以推广为

$$\begin{aligned}&\boldsymbol{M}(\boldsymbol{z}, x, t)\boldsymbol{z}_t + \boldsymbol{K}(\boldsymbol{z}, x, t)\boldsymbol{z}_x\\ &\quad = \nabla_{\boldsymbol{z}} B(\boldsymbol{z}, x, t) + \frac{\partial \boldsymbol{F}(\boldsymbol{z}, x, t)}{\partial t} + \frac{\partial \boldsymbol{G}(\boldsymbol{z}, x, t)}{\partial x},\end{aligned} \tag{6.1.7}$$

其中 $\boldsymbol{F} = \boldsymbol{F}(\boldsymbol{z}, x, t)$ 和 $\boldsymbol{G} = \boldsymbol{G}(\boldsymbol{z}, x, t)$ 是两个向量函数. 上面的系统要对称, 它的结构矩阵和势函数必须满足一定的对称条件. 这个对称条件事实上就等价于第四章中 Marsden 给出的外力作用垂直于群作用条件. 这两种对称条件从形式上看之所以不同, 是因为我们是在一般辛结构下讨论对称, 而 Marsden 等人是在标准辛结构下讨论受外力作用的

系统的对称条件. 对于 $\boldsymbol{M}$ 和 $\boldsymbol{F}$, $\boldsymbol{K}$ 和 $\boldsymbol{G}$ 的关系, 我们在下一节具体给出.

结构 (6.1.7) 之所以有讨论意义, 有两个方面原因: 第一, 在形式上它是多辛 Hamilton 系统 (6.1.1) 的自然推广, 同时也是从常微分方程的 Birkhoff 表示到偏微分方程的推广; 第二, 我们希望通过这个形式来研究受外力作用的不守恒系统的几何结构, 跟第五章的结果做一个比较. 下面我们先看一个具体的例子, 以便对这种系统有一个直观的认识. 考虑方程

$$u_{tt} - u_{xx} + u + \alpha u_t + \beta u_x = 0. \tag{6.1.8}$$

这个方程是描述空气中带磨擦的线性弦振动方程, 其中的耗散项是 u_t 和 u_x. 我们引进新变量 $p = u_t$ 和 $q = u_x$, 方程 (6.1.8) 变为一个一阶方程组

$$\begin{aligned} &u_t = p, \\ &u_x = q, \\ &p_t - q_x + u + \alpha p + \beta q = 0, \end{aligned} \tag{6.1.9}$$

或者令 $\boldsymbol{z} = (u, p, q)^{\mathrm{T}} \in \mathbf{R}^3$, 则它可以表示为系统 (6.1.7) 的形式, 其中

$$\boldsymbol{M} = \begin{pmatrix} 0 & \mathrm{e}^{\alpha t-\beta x} & 0 \\ -\mathrm{e}^{\alpha t-\beta x} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{K} = \begin{pmatrix} 0 & 0 & -\mathrm{e}^{\alpha t-\beta x} \\ 0 & 0 & 0 \\ \mathrm{e}^{\alpha t-\beta x} & 0 & 0 \end{pmatrix}, \tag{6.1.10}$$

势函数为

$$\begin{aligned} &\boldsymbol{F} = \left(-\frac{1}{2}\mathrm{e}^{\alpha t-\beta x}p, \frac{1}{2}\mathrm{e}^{\alpha t-\beta x}u, 0\right)^{\mathrm{T}}, \\ &\boldsymbol{G} = \left(\frac{1}{2}\mathrm{e}^{\alpha t-\beta x}q, 0, -\frac{1}{2}\mathrm{e}^{\alpha t-\beta x}u\right)^{\mathrm{T}}, \\ &B = -\frac{1}{2}\mathrm{e}^{\alpha t-\beta x}(u^2 + p^2 - q^2 + \alpha up + \beta uq). \end{aligned} \tag{6.1.11}$$

§6.2 Birkhoff 多辛结构

有了上一节的介绍，我们就可以给出多辛 Birkhoff 系统的定义了. 通过后面的几个命题和定理，我们会对这样的定义有进一步理解.

定义 6.2.1 如下形式的方程组称为**多辛 Birkhoff 系统**或 **Birkhoff 形式**:

$$\left(\frac{\partial F_\nu(\boldsymbol{z},x,t)}{\partial z^\mu}-\frac{\partial F_\mu(\boldsymbol{z},x,t)}{\partial z^\nu}\right)\frac{\partial z_\nu}{\partial t}+\left(\frac{\partial G_\nu(\boldsymbol{z},x,t)}{\partial z^\mu}-\frac{\partial G_\mu(\boldsymbol{z},x,t)}{\partial z^\nu}\right)\frac{\partial z_\nu}{\partial x}$$
$$-\left(\frac{\partial B(\boldsymbol{z},x,t)}{\partial z^\mu}+\frac{\partial F_\mu(\boldsymbol{z},x,t)}{\partial t}+\frac{\partial G_\mu(\boldsymbol{z},x,t)}{\partial x}\right)=0,$$
$$\mu,\nu=1,\cdots,n, \tag{6.2.1}$$

其中标量函数 $B(\boldsymbol{z},x,t)$ 及向量函数 $\boldsymbol{F}(\boldsymbol{z},x,t)$ 和 $\boldsymbol{G}(\boldsymbol{z},x,t)$ 统称为 **Birkhoff 函数**, 特别 B 称为 **Birkhoff 因子**.

一个偏微分方程什么时候可以表示成多辛 Birkhoff 形式? 理论上讲, 任何一个非奇异的偏微分方程都可以表示成多辛 Birkhoff 形式, 但是具体实现比较困难. 现在我们考虑一个一般的一阶 Newton 系统:

$$N^\mu(\boldsymbol{z})=M_{\mu\nu}(\boldsymbol{z},x,t)z_t^\nu+K_{\mu\nu}(\boldsymbol{z},x,t)z_x^\nu+D_\mu(\boldsymbol{z},x,t)=0,$$
$$\mu=1,\cdots,n. \tag{6.2.2}$$

下面给出它的变分对称性. 我们在第三章已经详细阐述了变分对称性的定义及相关结果, 这里应用这些结果证明下面的定理.

定理 6.2.1 协变一阶方程 (6.2.1) 对称的充分必要条件是, 存在构形空间中点 $(\boldsymbol{z},x,t)$ 的星形区域 $\widetilde{\mathfrak{R}}^*$, 在这个区域里, 下述条件成立:

$$\left.\begin{aligned}&M_{\mu\nu}+M_{\nu\mu}=0,\\&K_{\mu\nu}+K_{\nu\mu}=0,\\&\frac{\partial M_{\mu\nu}}{\partial z^\tau}+\frac{\partial M_{\nu\tau}}{\partial z^\mu}+\frac{\partial M_{\tau\mu}}{\partial z^\nu}=0,\\&\frac{\partial K_{\mu\nu}}{\partial z^\tau}+\frac{\partial K_{\nu\tau}}{\partial z^\mu}+\frac{\partial K_{\tau\mu}}{\partial z^\nu}=0,\\&\frac{\partial M_{\mu\nu}}{\partial t}+\frac{\partial K_{\mu\nu}}{\partial x}=\frac{\partial D_\mu}{\partial z^\nu}-\frac{\partial D_\nu}{\partial z^\mu},\end{aligned}\right\}\quad \mu,\nu=1,\cdots,n. \tag{6.2.3}$$

证明 假设 $\boldsymbol{N}$ 存在连续的 Frechét 导数, 即

$$N_{\nu}^{\mu'}\phi = \lim_{\varepsilon\to 0}\frac{N^{\mu}(z^{\nu}+\varepsilon\phi)-N^{\mu}(z^{\nu})}{\varepsilon}, \quad \mu,\nu=1,\cdots,n, \tag{6.2.4}$$

这里 ϕ 是从 $\mathbf{R}^2$ 到 $\mathbf{R}$ 的任意函数, 那么方程组 (6.2.2) 是对称的, 当且仅当 $\boldsymbol{N}$ 的 Frechét 导算子 $N_{\nu}^{\mu'}$ 是自共轭算子, 即

$$N_{\nu}^{\mu'} = N_{\mu}^{\nu'}. \tag{6.2.5}$$

很容易证明上面的等式等价于条件 (6.2.3). □

我们之所以讨论协变的一阶方程, 原因在于: 首先, 反变的一阶方程组假如是对称的, 那么它一定是多辛 Hamilton 系统, 我们不需要再特别讨论它; 其次, 协变的一阶方程组可以从一个最一般的一阶泛函变分导出. 这个泛函就是 Pfaff 作用泛函:

$$\pounds(\boldsymbol{z},x,t) = \iint_{(x,t)}[F_{\nu}(\boldsymbol{z},x,t)z_t^{\nu}+G_{\nu}(\boldsymbol{z},x,t)z_x^{\nu}-B(\boldsymbol{z},x,t)]\mathrm{d}t\mathrm{d}x. \tag{6.2.6}$$

定义泛函 (6.2.6) 中的被积函数 $L = F_{\nu}(\boldsymbol{z},x,t)z_t^{\nu} + G_{\nu}(\boldsymbol{z},x,t)z_x^{\nu} - B(\boldsymbol{z},x,t)$ 为 Lagrange 函数. 一阶的方程组不存在共轭 Legendre 变换, 所以一个对称方程组要么本身就是多辛 Hamilton 方程, 要么是多辛 Birkhoff 方程.

命题 6.2.1 (Birkhoff 系统的对称性) 协变一阶系统 (6.2.2) 对称的充分必要条件是, 在空间 $\mathbf{R}^n\times\mathbf{R}^2$ 的星形区域 $\widetilde{\mathfrak{R}}^*$ 里, 它可以表示成 Birkhoff 形式 (6.2.1), 即

$$\begin{aligned}
&M_{\mu\nu}(\boldsymbol{z},x,t)z_t^{\nu} + K_{\mu\nu}(\boldsymbol{z},x,t)z_x^{\nu} + D_{\mu}(\boldsymbol{z},x,t)\\
&= \left(\frac{\partial F_{\nu}(\boldsymbol{z},x,t)}{\partial z^{\mu}} - \frac{\partial F_{\mu}(\boldsymbol{z},x,t)}{\partial z^{\nu}}\right)\frac{\partial z^{\nu}}{\partial t}\\
&\quad + \left(\frac{\partial G_{\nu}(\boldsymbol{z},x,t)}{\partial z^{\mu}} - \frac{\partial G_{\mu}(\boldsymbol{z},x,t)}{\partial z^{\nu}}\right)\frac{\partial z^{\nu}}{\partial x}\\
&\quad - \left(\frac{\partial B(\boldsymbol{z},x,t)}{\partial z^{\mu}} + \frac{\partial F_{\mu}(\boldsymbol{z},x,t)}{\partial t} + \frac{\partial G_{\mu}(\boldsymbol{z},x,t)}{\partial x}\right) = 0,\\
&\qquad \mu,\nu = 1,\cdots,n.
\end{aligned} \tag{6.2.7}$$

证明　定义一个 3-形式

$$\Omega = M_{\mu\nu}\mathrm{d}z^\mu \wedge \mathrm{d}z^\nu \wedge \mathrm{d}x - K_{\mu\nu}\mathrm{d}z^\mu \wedge \mathrm{d}z^\nu \wedge \mathrm{d}t + 2D_\mu \mathrm{d}z^\mu \wedge \mathrm{d}t \wedge \mathrm{d}x. \tag{6.2.8}$$

充分性　如果协变一阶系统可以表示成 Birkhoff 形式, 那么意味着以上 3-形式是一个闭形式, 即

$$\mathrm{d}\Omega = 0. \tag{6.2.9}$$

Ω 为闭形式的可积条件就是对称条件 (6.2.3).

必要性　若协变一阶系统 (6.2.2) 对称, 则对称条件 (6.2.3) 成立, 从而 Ω 为闭的. 那么, 根据 Poincaré 引理的逆命题知道, 形式 Ω 在星形区域 $\widetilde{\mathfrak{R}}^*$ 上是确切的, 也就是存在一个 2-形式 Θ, 使得

$$\Omega = \mathrm{d}\Theta, \tag{6.2.10}$$

也即存在 Birkhoff 函数 $\boldsymbol{F}(\boldsymbol{z},x,t)$, $\boldsymbol{G}(\boldsymbol{z},x,t)$ 和 $B(\boldsymbol{z},x,t)$, 使得

$$\Theta = F_\mu \mathrm{d}z^\mu \wedge \mathrm{d}x - G_\mu \mathrm{d}z^\mu \wedge \mathrm{d}t - B\mathrm{d}t \wedge \mathrm{d}x. \tag{6.2.11}$$

根据 (6.2.10) 式, 方程 (6.2.2) 一定具有形式 (6.2.7).　□

3-形式 Ω 是 Poincaré-Cartan 形式最一般的推广, 和我们在第四章中通过场论变分得到的 3-形式是一致的. Poincaré-Cartan 形式在 Bridges[Bri97a,BR01,Bri97b] 和 Marsden 等人的理论中都有所说明.

§6.3　Birkhoff 多辛守恒律和多辛格式

多辛 Birkhoff 系统 (6.2.7) 也定义了两个 2-形式:

$$\omega = \boldsymbol{M}(\boldsymbol{z},x,t)\mathrm{d}\boldsymbol{z} \wedge \mathrm{d}\boldsymbol{z},$$
$$\kappa = \boldsymbol{K}(\boldsymbol{z},x,t)\mathrm{d}\boldsymbol{z} \wedge \mathrm{d}\boldsymbol{z}.$$

我们对原方程组 (6.2.7) 两边作外微分, 再外积 $\mathrm{d}\boldsymbol{z}$, 可以得到

$$\begin{aligned}0 &= \mathrm{d}(\boldsymbol{M}\boldsymbol{z}_t) \wedge \mathrm{d}\boldsymbol{z} + \mathrm{d}(\boldsymbol{K}\boldsymbol{z}_x) \wedge \mathrm{d}\boldsymbol{z} + \mathrm{d}\boldsymbol{D} \wedge \mathrm{d}\boldsymbol{z}\\ &= M_{\mu\nu}\mathrm{d}z_t^\nu \wedge \mathrm{d}z^\mu + K_{\mu\nu}\mathrm{d}z_x^\nu \wedge \mathrm{d}z^\mu + \frac{\partial M_{\mu\nu}}{\partial z^\beta} z_t^\nu \mathrm{d}z^\beta \wedge \mathrm{d}z^\mu\end{aligned}$$

$$+\frac{\partial K_{\mu\nu}}{\partial z^\beta}z_x^\nu \mathrm{d}z^\beta \wedge \mathrm{d}z^\mu + \nabla_{\boldsymbol{zz}}B\mathrm{d}\boldsymbol{z}\wedge\mathrm{d}\boldsymbol{z} + \frac{\partial^2 F_\mu}{\partial z^\beta \partial t}\mathrm{d}z^\beta\wedge\mathrm{d}z^\mu$$
$$+\frac{\partial^2 G_\mu}{\partial z^\beta\partial x}\mathrm{d}z^\beta\wedge\mathrm{d}z^\mu. \tag{6.3.1}$$

$\nabla_{\boldsymbol{zz}}B$ 是对称的, 利用对称条件 (6.2.3), 得到

$$\boldsymbol{0} = \frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{M}\mathrm{d}\boldsymbol{z}\wedge\mathrm{d}\boldsymbol{z}) + \frac{\mathrm{d}}{\mathrm{d}x}(\boldsymbol{K}\mathrm{d}\boldsymbol{z}\wedge\mathrm{d}\boldsymbol{z}), \tag{6.3.2}$$

也就是多辛守恒律

$$\frac{\mathrm{d}\omega}{\mathrm{d}t} + \frac{\mathrm{d}\kappa}{\mathrm{d}x} = 0.$$

我们称这个守恒律为 **Birkhoff 多辛守恒律**. 它看起来和 Hamilton 多辛守恒律 (6.1.3) 完全一样, 但是由于系数矩阵 $\boldsymbol{M}_t \neq \boldsymbol{0}$ 和 $\boldsymbol{K}_x \neq \boldsymbol{0}$, 所以它们是和前者不完全一样的, 后者实质上更像是一种耗散规律, 在后面有具体的例子更能生动地说明这一点. 另外, 虽然多辛守恒和形式能量或动量守恒不一样, 但是在 Bridges 的理论框架中, 他导出形式能量和动量守恒律的路径与导出多辛守恒律的路径是一样的, 只是两者考虑的几何空间有差异. 所以笔者认为, 假如在完全一样的空间上考虑, 多辛守恒与形式能量和动量守恒是等价的. 但 Birkhoff 系统不能像自治的 Hamilton 系统一样具有能量和动量守恒律, 从这一点也说明了 (6.3.2) 式只是形式的守恒律, 事实上它反映的是系统的耗散规律. 那么, 我们能否像得出多辛守恒律一样也得出反映 Birkhoff 系统耗散规律的形式能量和动量守恒律呢?

我们在这里先讨论一下一般多辛结构下自治的 Hamilton 系统的形式能量和动量守恒律. 一个形如

$$\boldsymbol{M}(\boldsymbol{z})\boldsymbol{z}_t + \boldsymbol{K}(\boldsymbol{z})\boldsymbol{z}_x = \nabla_{\boldsymbol{z}}S(\boldsymbol{z}) \tag{6.3.3}$$

的对称系统就是一般的多辛 Hamilton 系统. 因为它是对称的, 所以显然其中的系数矩阵 $\boldsymbol{M}$ 和 $\boldsymbol{K}$ 定义的 2-形式在某个星形区域上也是闭的, 即存在向量函数 $\boldsymbol{F}(\boldsymbol{z})$ 和 $\boldsymbol{G}(\boldsymbol{z})$, 使得

$$M_{\mu\nu} = \frac{\partial F_\nu}{\partial z^\mu} - \frac{\partial F_\mu}{\partial z^\nu},\quad K_{\mu\nu} = \frac{\partial G_\nu}{\partial z^\mu} - \frac{\partial G_\mu}{\partial z^\nu},$$
$$\mu,\nu = 1,\cdots,n. \tag{6.3.4}$$

那么, 由原方程我们可以定义形式能量和动量守恒律, 例如

$$\partial_t\mathcal{E}(\boldsymbol{z})+\partial_x\mathcal{F}(\boldsymbol{z})=0 \quad 和 \quad \partial_x\mathcal{I}(\boldsymbol{z})+\partial_t\mathcal{G}(\boldsymbol{z})=0, \tag{6.3.5}$$

其中

$$\begin{aligned}&\mathcal{E}(\boldsymbol{z})=S(\boldsymbol{z})-\frac{1}{2}F_\mu z_x^\mu, \quad \mathcal{F}(\boldsymbol{z})=\frac{1}{2}F_\mu z_t^\mu,\\&\mathcal{I}(\boldsymbol{z})=S(\boldsymbol{z})-\frac{1}{2}G_\mu z_t^\mu, \quad \mathcal{G}(\boldsymbol{z})=\frac{1}{2}G_\mu z_x^\mu.\end{aligned} \tag{6.3.6}$$

这样的局部的形式能量和动量守恒加上适当的边界条件可以导出整体的形式能量和动量守恒.

现在我们再回过头来讨论 Birkhoff 系统 (6.2.1). 如果我们引进新的变量 s_1, s_2 和 $t_1=t$, $x_1=x$, 那么 Birkhoff 系统 (6.2.1) 可以被一个新的系统所包含, 这个新系统是一般辛结构下非自治的 Hamilton 系统, 表示为

$$\begin{aligned}&\begin{pmatrix}0&0&1&0&\mathbf{0}\\0&0&0&0&\mathbf{0}\\-1&0&0&0&\boldsymbol{\Delta}_1^{\mathrm{T}}\\0&0&0&0&\mathbf{0}\\\mathbf{0}&\mathbf{0}&-\boldsymbol{\Delta}_1&\mathbf{0}&\boldsymbol{M}\end{pmatrix}\partial_t\begin{pmatrix}s_1\\s_2\\t_1\\x_1\\\boldsymbol{z}\end{pmatrix}\\&\quad+\begin{pmatrix}0&0&0&0&\mathbf{0}\\0&0&0&1&\mathbf{0}\\0&0&0&0&\mathbf{0}\\0&-1&0&0&\boldsymbol{\Delta}_2^{\mathrm{T}}\\\mathbf{0}&\mathbf{0}&\mathbf{0}&-\boldsymbol{\Delta}_2&\boldsymbol{K}\end{pmatrix}\partial_x\begin{pmatrix}s_1\\s_2\\t_1\\x_1\\\boldsymbol{z}\end{pmatrix}\\&=\begin{pmatrix}1\\1\\0\\0\\-\boldsymbol{D}(\boldsymbol{z},t,x)-\boldsymbol{F}_t(\boldsymbol{z},x,t)-\boldsymbol{G}_x(\boldsymbol{z},x,t)\end{pmatrix},\end{aligned} \tag{6.3.7}$$

这里 $\boldsymbol{\Delta}_1 = \dfrac{\partial \boldsymbol{F}(\boldsymbol{z}, t_1, x_1)}{\partial t_1}$, $\boldsymbol{\Delta}_2 = \dfrac{\partial \boldsymbol{G}(\boldsymbol{z}, t_1, x_1)}{\partial x_1}$. 这相当于在方程组 (6.3.7) 的左边把原来函数中的自变量 x 和 t 都替换为因变量 x_1 和 t_1, 而在右边函数中的自变量不替换, 那么很显然它的右边是一个梯度函数. 我们把这个新系统简化表示为

$$\widetilde{\boldsymbol{M}}(\widetilde{z})\widetilde{z}_t + \widetilde{\boldsymbol{K}}(\widetilde{z})\widetilde{z}_x = \boldsymbol{\nabla}_{\widetilde{z}}\widetilde{S}(\widetilde{z}, x, t), \tag{6.3.8}$$

其中 $\widetilde{\boldsymbol{z}} = (s_1, s_2, t_1, x_1, \boldsymbol{z}^{\mathrm{T}})^{\mathrm{T}}$, $\widetilde{S} = s_1 + s_2 + B(\boldsymbol{z}, x, t)$.

对非自治的多辛 Hamilton 系统 (6.3.8) 引进四个新变量 s_3, s_4 和 $x_2 = x$, $t_2 = t$, 那么它可以被一个自治的 Hamilton 系统所包含, 这个自治的 Hamilton 系统表示为

$$\overline{\boldsymbol{M}}(\overline{z})\overline{z}_t + \overline{\boldsymbol{K}}(\overline{z})\overline{z}_x = \boldsymbol{\nabla}_{\overline{z}}\overline{S}(\overline{z}), \tag{6.3.9}$$

其中

$$\overline{\boldsymbol{M}}(\overline{z}) = \begin{pmatrix} 0 & 0 & 1 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & \mathbf{0} \\ -1 & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \widetilde{\boldsymbol{M}} \end{pmatrix}, \quad \overline{\boldsymbol{K}}(\overline{z}) = \begin{pmatrix} 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & 1 & \mathbf{0} \\ 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & -1 & 0 & 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \widetilde{\boldsymbol{K}}, \end{pmatrix}, \tag{6.3.10}$$

$$\begin{aligned} &\nabla_{\overline{z}}\overline{S} = \left(1, 1, \frac{\partial}{\partial t_2}B(\boldsymbol{z}, t_2, x_2), \frac{\partial}{\partial x_2}B(\boldsymbol{z}, t_2, x_2), 1, 1, 0, 0, \frac{\partial}{\partial \boldsymbol{z}}B(\boldsymbol{z}, t_2, x_2)\right)^{\mathrm{T}}, \\ &\overline{\boldsymbol{z}} = (s_3, s_4, t_2, x_2, s_1, s_2, t_1, x_1, \boldsymbol{z}^{\mathrm{T}})^{\mathrm{T}}, \\ &\overline{S}(\overline{\boldsymbol{z}}) = B(\boldsymbol{z}, x_2, t_2,) + s_1 + s_2 + s_3 + s_4. \end{aligned} \tag{6.3.11}$$

这样利用一般辛结构下的 Hamilton 能量和动量守恒律 (6.3.5) 和 (6.3.6), 我们就得到了 Birkhoff 系统的形式能量和动量守恒律. 具体公式的计算很简单, 但步骤很多, 本章不再详细讨论.

在数值计算方面, 我们也希望逼近 Birkhoff 系统 (6.2.1) 的数值格式能够保持它的相应离散的 Birkhoff 多辛守恒律. 假设已知系统 (6.2.1)

的一个离散形式

$$\begin{aligned}&\boldsymbol{M}_{i,j}\partial_t^{i,j}\boldsymbol{z}_{i,j}+\boldsymbol{K}_{i,j}\partial_x^{i,j}\boldsymbol{z}_{i,j}\\&\quad=(\nabla_{\boldsymbol{z}}B(\boldsymbol{z}_{i,j},x_i,t_j))_{i,j}+(\partial_t\boldsymbol{F}(\boldsymbol{z}_{i,j},x_i,t_j))_{i,j}\\&\qquad+(\partial_x\boldsymbol{G}(\boldsymbol{z}_{i,j},x_i,t_j))_{i,j},\quad \boldsymbol{z}_{i,j}=\boldsymbol{z}(x_i,t_j),\end{aligned}\tag{6.3.12}$$

对应的离散多辛守恒律为

$$\partial_t^{i,j}\omega_{i,j}+\partial_x^{i,j}\kappa_{i,j}=\boldsymbol{0},\tag{6.3.13}$$

这里

$$\omega_{i,j}=\langle\boldsymbol{M}_{i,j}U_{i,j},V_{i,j}\rangle,\quad \kappa_{i,j}=\langle\boldsymbol{K}_{i,j}U_{i,j},V_{i,j}\rangle.\tag{6.3.14}$$

定义 6.3.1 如果离散方程 (6.3.12) 有形如 (6.3.13) 式的 Birkhoff 多辛守恒律, 那么它决定的离散格式称为系统 (6.2.1) 的一个 **Birkhoff 多辛格式**.

§6.4　线性阻尼振动方程的 Birkhoff 形式

不管有没有受到外力作用或者说有没有耗散项, 多辛 Hamilton 系统和多辛 Birkhoff 系统的本质区别在于它们的多辛结构. 不依赖于任何变量的多辛结构下的 Hamilton 系统的多辛几何算法已经发展得很完善了, 而依赖于函数变量的一般 Hamilton 多辛结构, 包含了系统较复杂的能量或者其他信息, 因此其相应的多辛几何算法研究较复杂. Birkhoff 多辛结构较之一般 Hamilton 多辛结构更复杂, 它除了依赖于函数变量还依赖于自变量, 依赖于自变量使得它还包含了对系统作用的外力的信息, 从而使得构造 Birkhoff 多辛格式要比构造 Hamilton 多辛格式复杂得多. 下面我们以弦振动方程为例, 构造一个 Birkhoff 多辛格式.

考虑弦振动方程

$$u_{tt}-u_{xx}+u+2u_t=0,\tag{6.4.1}$$

其中耗散项 $2u_t$ 代表了不对称的外力. 引进两个势函数 p 和 q, 使得 $u_t=p$, $u_x=q$, 则上面的方程就等价于下面的方程组:

$$\begin{aligned} &u_t=p,\\ &u_x=q,\\ &p_t-q_x+u+2p=0. \end{aligned} \tag{6.4.2}$$

在方程组 (6.4.2) 各方程的两边同时乘以 e^{2t}, 我们就可以得到原来方程的 Birkhoff 表示

$$\begin{aligned} &\mathrm{e}^{2t}u_t=\mathrm{e}^{2t}p,\\ &\mathrm{e}^{2t}u_x=\mathrm{e}^{2t}q,\\ &\mathrm{e}^{2t}p_t-\mathrm{e}^{2t}q_x+\mathrm{e}^{2t}u+2\mathrm{e}^{2t}p=0. \end{aligned} \tag{6.4.3}$$

方程组 (6.4.3) 显然满足对称条件. 它有一个依赖于时间变量 t 的多辛 Birkhoff 形式

$$\begin{aligned} &\begin{pmatrix} 0 & \mathrm{e}^{2t} & 0\\ -\mathrm{e}^{2t} & 0 & 0\\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} u_t\\ p_t\\ q_t \end{pmatrix}+\begin{pmatrix} 0 & 0 & -\mathrm{e}^{2t}\\ 0 & 0 & 0\\ \mathrm{e}^{2t} & 0 & 0 \end{pmatrix}\begin{pmatrix} u_x\\ p_x\\ q_x \end{pmatrix}\\ &=\begin{pmatrix} -\mathrm{e}^{2t}u-2\mathrm{e}^{2t}p\\ -\mathrm{e}^{2t}p\\ \mathrm{e}^{2t}q \end{pmatrix}, \end{aligned} \tag{6.4.4}$$

相应的 Birkhoff 函数是

$$\begin{aligned} &\boldsymbol{F}=\left(-\frac{1}{2}\mathrm{e}^{2t}p,\frac{1}{2}\mathrm{e}^{2t}u,0\right)^{\mathrm{T}},\\ &\boldsymbol{G}=\left(\frac{1}{2}\mathrm{e}^{2t}q,0,-\frac{1}{2}\mathrm{e}^{2t}u\right)^{\mathrm{T}},\\ &B=-\frac{1}{2}\mathrm{e}^{2t}(u^2+p^2-q^2+2up). \end{aligned} \tag{6.4.5}$$

从多辛 Birkhoff 形式 (6.4.4) 直接推导, 得到它的 Birkhoff 多辛守恒律

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{e}^{2t}\mathrm{d}p\wedge\mathrm{d}u)-\frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^{2t}\mathrm{d}q\wedge\mathrm{d}u)=0, \tag{6.4.6}$$

如果消去它的耗散因子 e^{2t} (也是系统多辛结构的关键元素), 上面的守恒律就不再守恒了, 变成如下形式:

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{d}p \wedge \mathrm{d}u) - \frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{d}q \wedge \mathrm{d}u) = -2\mathrm{d}p \wedge \mathrm{d}u. \tag{6.4.7}$$

如果对原来的振动方程引进变量 P 和 Q, 使得 $P = \mathrm{e}^{2t}u_t$, $Q = \mathrm{e}^{2t}u_x$, 那么原方程等价于多辛 Hamilton 形式

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \partial_t \begin{pmatrix} u \\ P \\ Q \end{pmatrix} + \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \partial_x \begin{pmatrix} u \\ P \\ Q \end{pmatrix} = \begin{pmatrix} -\mathrm{e}^{2t}u \\ -\mathrm{e}^{-2t}P \\ \mathrm{e}^{-2t}Q \end{pmatrix}, \tag{6.4.8}$$

其 Hamilton 函数为

$$H = -\frac{1}{2}\mathrm{e}^{2t}u^2 + \frac{1}{2}\mathrm{e}^{-2t}(Q^2 - P^2).$$

上面的形式有 Hamilton 多辛守恒律

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{d}P \wedge \mathrm{d}u) + \frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{d}u \wedge \mathrm{d}Q) = 0. \tag{6.4.9}$$

我们还可以对原来的振动方程引进不同表示的变量, 得到更多不同的多辛形式. 比如又一个多辛 Birkhoff 形式如下:

$$\begin{pmatrix} 0 & \mathrm{e}^{-2t} & 0 \\ -\mathrm{e}^{-2t} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \partial_t \begin{pmatrix} u \\ P \\ Q \end{pmatrix} + \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \partial_x \begin{pmatrix} u \\ P \\ Q \end{pmatrix} = \begin{pmatrix} -u \\ -\mathrm{e}^{-4t}P \\ Q \end{pmatrix}. \tag{6.4.10}$$

将多辛 Birkhoff 形式 (6.4.4) 和多辛 Hamilton 形式 (6.4.8) 进行比较. 后者是很严格简洁的表示, 在数值计算方面已经存在现成的多辛逼近, 比如 Preissman 格式、多辛 PRK 方法等. 然而后者的缺点是: 系统的变量并非实际期望值或观测值, 不一定有实际的物理意义. 本例中 P 和 H 已经不是习惯的能量和动量的期望值或观测值. 我们介绍 Birkhoff 多辛表示, 对解决这个问题也许会有所帮助. 首先多辛 Birkhoff 形式 (6.4.4) 的未知函数为 $\boldsymbol{z} = (u, p, q)$, 其中 u 代表系统 (6.4.1) 的观测值,

p 是它的线性动量. 其次 Birkhoff 形式在 t 方向的辛张量

$$\boldsymbol{M}(\boldsymbol{z},t,x)=\begin{pmatrix}0 & \mathrm{e}^{2t} & 0\\ -\mathrm{e}^{2t} & 0 & 0\\ 0 & 0 & 0\end{pmatrix} \tag{6.4.11}$$

代表了不守恒、不对称的外力 $f=-2u_t$. 还有 Birkhoff 多辛守恒律 (6.4.6) 带有系统因为外力而产生的耗散因子 e^{2t}, 它在变换

$$P=\mathrm{e}^{2t}p,\quad Q=\mathrm{e}^{2t}q \tag{6.4.12}$$

下, 和 Hamilton 多辛守恒律 (6.4.9) 是等价的, 但它本身不能算是纯粹的守恒律, 所以说它所在的 Birkhoff 框架也更能代表耗散系统. 很自然地, 我们认为由此框架出发构造的数值格式能更好地保持原系统的几何结构.

下面我们对 Birkhoff 方程 (6.4.1) 进行离散, 得到

$$\begin{aligned}&\frac{\mathrm{e}^{\delta t}p_{i+1/2,j+1}-\mathrm{e}^{-\delta t}p_{i+1/2,j}}{\delta t}+\frac{q_{i+1,j+1/2}-q_{i,j+1/2}}{\delta x}\\&\quad=-u_{i+1/2,j+1/2},\\&\frac{u_{i+1/2,j+1}-u_{i+1/2,j}}{\delta t}=p_{i+1/2,j+1/2},\\&\frac{u_{i+1,j+1/2}-u_{i,j+1/2}}{\delta x}=q_{i+1/2,j+1/2}.\end{aligned} \tag{6.4.13}$$

因为系统的 Birkhoff 多辛结构只依赖于时间变量 t, 而不依赖于空间变量 x, 所以 ∂_x 的离散形式我们一律采用中点离散, 而 ∂_t 的离散就稍有不同, 在离散点 $(x_{i+1/2},t_{j+1/2})$ 处分别是

$$\begin{aligned}\mathrm{e}^{2t}u_t(i+1/2,j+1/2)&=\frac{\mathrm{e}^{2t_{j+1/2}}u_{i+1/2,j+1}-\mathrm{e}^{2t_{j+1/2}}u_{i+1/2,j}}{\delta t},\\ \mathrm{e}^{2t}q_t(i+1/2,j+1/2)&=\frac{\mathrm{e}^{2t_{j+1/2}}q_{i+1/2,j+1}-\mathrm{e}^{2t_{j+1/2}}q_{i+1/2,j}}{\delta t},\\ \mathrm{e}^{2t}p_t(i+1/2,j+1/2)&=\frac{\mathrm{e}^{2t_{j+1}}p_{i+1/2,j+1}-\mathrm{e}^{2t_j}p_{i+1/2,j}}{\delta t},\end{aligned} \tag{6.4.14}$$

另外, 它们的中点分解成

$$
\begin{aligned}
p_{i+1/2,j+1/2} &= \frac{1}{4}\mathrm{e}^{\delta t}(p_{i+1,j+1}+p_{i,j+1})+\frac{1}{4}\mathrm{e}^{-\delta t}(p_{i+1,j}+p_{i,j}),\\
u_{i+1/2,j+1/2} &= \frac{1}{4}(u_{i+1,j+1}+u_{i,j+1}+u_{i+1,j}+u_{i,j}),\\
q_{i+1/2,j+1/2} &= \frac{1}{4}(q_{i+1,j+1}+q_{i,j+1}+q_{i+1,j}+q_{i,j}).
\end{aligned}
\tag{6.4.15}
$$

用同样的格式离散 Birkhoff 多辛守恒律 (6.4.6), 得到一个离散空间上的多辛守恒律

$$
\begin{aligned}
&\frac{\mathrm{e}^{\delta t}\mathrm{d}p_{i+1/2,j+1}\wedge \mathrm{d}u_{i+1/2,j+1}-\mathrm{e}^{-\delta t}\mathrm{d}p_{i+1/2,j}\wedge \mathrm{d}u_{i+1/2,j}}{\delta t}\\
&\quad -\frac{\mathrm{d}q_{i+1,j+1/2}\wedge \mathrm{d}u_{i+1,j+1/2}-\mathrm{d}q_{i,j+1/2}\wedge \mathrm{d}u_{i,j+1/2}}{\delta x}=0.
\end{aligned}
\tag{6.4.16}
$$

我们对离散方程 (6.4.13) 也作外微分, 再外乘离散空间上的 1-形式可以得到它的一个多辛守恒律, 这个守恒律和 Birkhoff 多辛守恒律 (6.4.6) 的离散形式 (6.4.16) 一致. 所以由离散等式 (6.4.14) 和 (6.4.15) 确定的格式是多辛 Birkhoff 系统 (6.4.1) 的一个多辛积分子. 下面进行数值实验.

我们在区间 $-\pi\leqslant x\leqslant\pi$ 上及 $t\geqslant 0$ 时考虑方程 (6.4.1) 的初边值解. 假如方程满足初值 $u(x,0)=\sin x$ 及周期边界条件, 则此方程有解析解

$$
u(x,t)=\exp(-t)\sin(x+t),
$$

且它有衰减的能量

$$
\int_{-\pi}^{\pi}(u_t+u)^2+u_x^2\mathrm{d}x=\pi\mathrm{e}^{2t}.
$$

分别用多辛格式 (6.4.13)—(6.4.15) 和下列中心格式数值模拟弦振动方程 (6.4.1):

$$
\begin{aligned}
&\frac{u_{i,j+1}-2u_{i,j}+u_{i,j-1}}{(\delta\tau)^2}-\frac{u_{i+1,j}-2u_{i,j}+u_{i-1,j}}{(\delta x)^2}\\
&\quad +u_{i,j}+2\frac{u_{i,j+1}-u_{i,j-1}}{2\delta t}=0,
\end{aligned}
$$

数值结果如图 6.4.1 和图 6.4.2 所示.

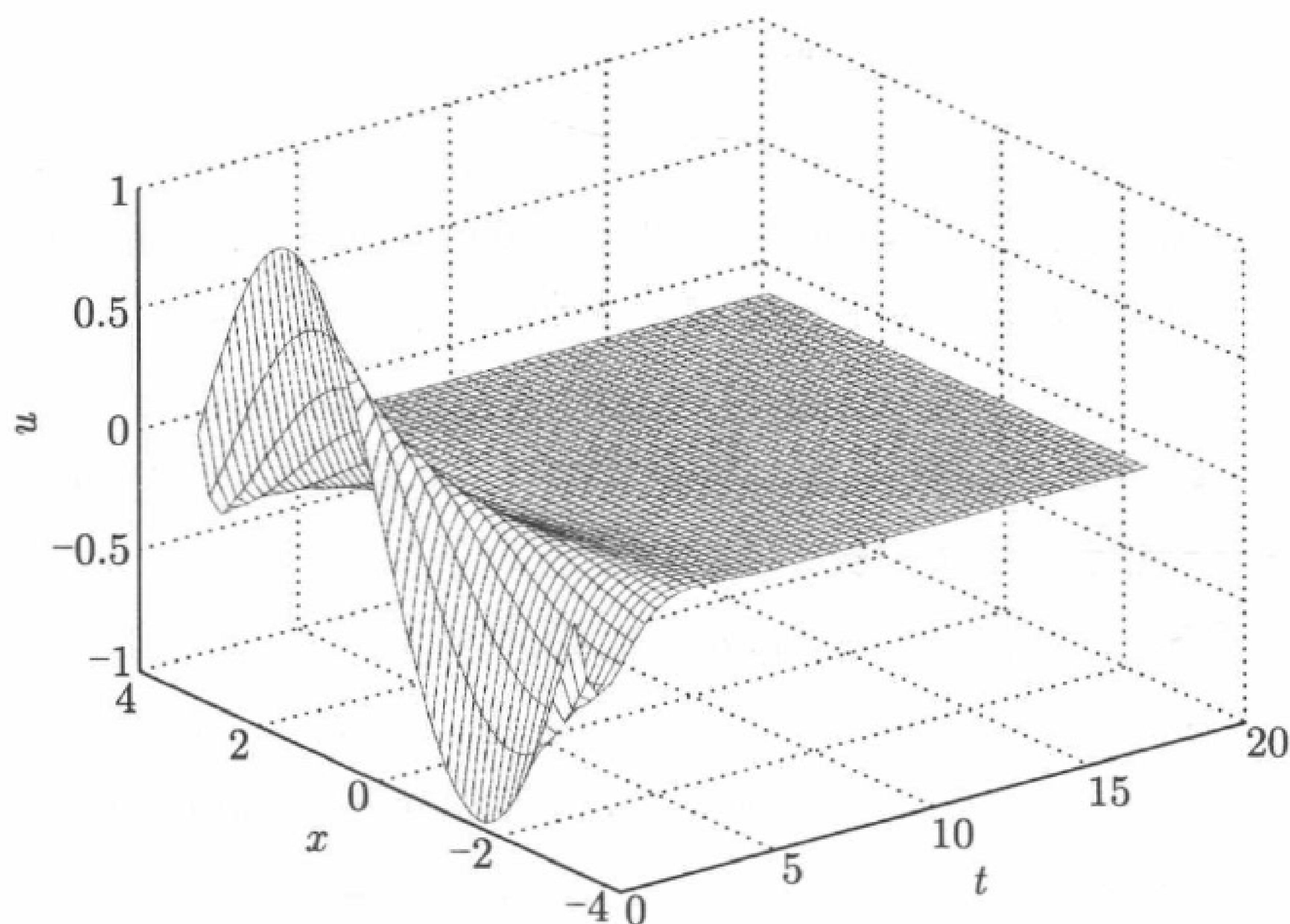


图 6.4.1 Birkhoff 多辛格式 (6.4.13)—(6.4.15) 的计算结果

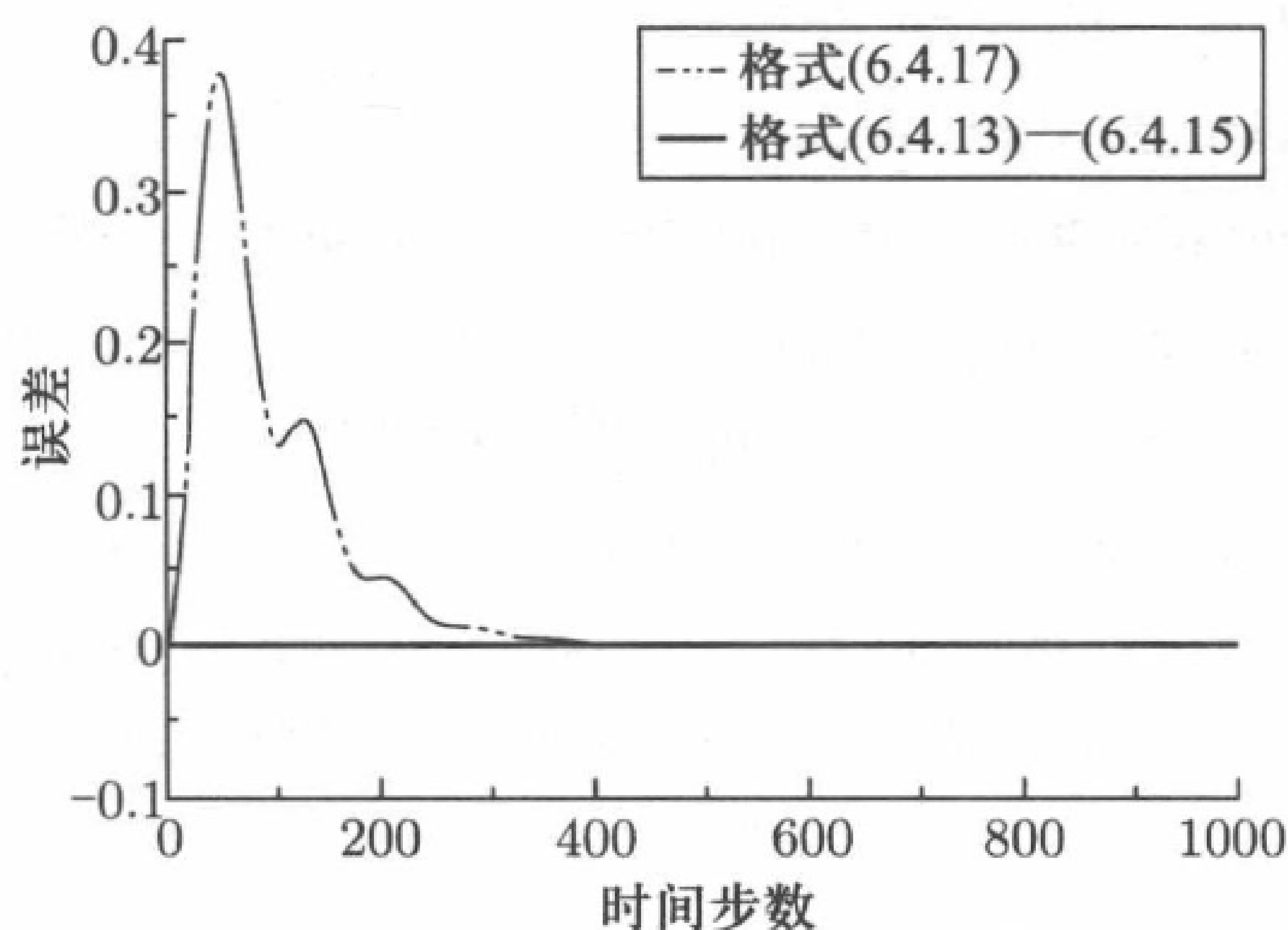


图 6.4.2 Birkhoff 多辛格式 (6.4.13)—(6.4.15) 和中心格式 (6.4.17) 的数值误差

图 6.4.1 表示多辛格式 (6.4.13)—(6.4.15) 在 $0 \leqslant t \leqslant 18$, 时间步长 $\delta t = 0.0311$, 空间步长 $\delta x = 0.0222$ 时的计算结果.

在时间 $t = j\delta t$ 处的误差定义为

$$\text{error} = \max_i |u_{i,j} - u(i\delta x, j\delta t)|.$$

图 6.4.2 给出了 Birkhoff 多辛格式 (6.4.13)—(6.4.15) 和中心格式

(6.4.17) 从 $t = 0$ 到 $t = 0.0311 \times 1000$ 的误差. 前者明显优于后者.

参考文献

[Bri97a] Bridges T J. Multi-symplectic structures and wave propagation. *Math Proc Cam Phil Soc*, 1997, 121: 147–190.

[BR01] Bridges T J, Reich S. Multi-symplectic integrators: numerical schemes for Hamiltonian PDEs that conserve symplecticity. *Phys Lett A*, 2001, 284: 184–193.

[Bri97b] Bridges T J. A geometric formulation of the conservation of wave action and its implications for signature and the classification of instabilities. *Proceeding: Royal Society of London A*, 1997, 453: 1365–1395.

[MPS98] Marsden J E, Patrick G P, Shkoller S. Multisymplectic geometry, variational integrators, and nonlinear PDEs. *Commun Math Phys*, 1998, 199: 351–395.

[Rei00] Reich S. Multi-symplectic Runge-Kutta collocation methods for Hamiltonian wave equations. *J Comp Phys*, 2000, 157: 473–499.

[San83b] Santilli R M. *Foundations of Theoretical Mechanics II*. 2nd edition. New York: Springer-Verlag, 1983.

[SQ10] Su Hongling, Qin Mengzhao, Wang Yushun, Scherer Rudolf. Multi-symplectic Birkhoffian structure for PDEs with dissipation terms. *Phys Lett A*, 2010, 374: 2410–2416.

第七章 无限维 Birkhoff 系统和生成泛函方法

本章我们介绍无限维 Birkhoff 系统、无限维 Birkhoff 辛结构以及用来构造无限维辛格式的生成泛函方法. 因为无限维 Hamilton 系统是 Birkhoff 系统的一个特殊情形, 并且关于它的概念、性质和算法有大量文献可以参阅, 所以为了帮助读者更好地理解后面的 Birkhoff 系统的定义和性质, 我们先简单介绍一下有关无限维 Hamilton 系统的基本概念、性质和算法. 由于所涉及的内容庞杂, 无法详细给出推导和证明, 有兴趣的读者可以参考文献 [MarRar99,Arn89,FW91,FQ09].

我们从介绍定义和记号开始. 假设 M 是一个微分流形, $T_{\boldsymbol{z}}M$ 是它的切空间, $\Omega : T_{\boldsymbol{z}}M \times T_{\boldsymbol{z}}M \to \mathbf{R}$ 是 M 上的一个非退化的闭 2-形式. 我们用 (M, Ω) 表示一个辛流形, $X : T_{\boldsymbol{z}}M \to M$ 表示一个向量场. 如果存在一个泛函 $\mathcal{H} : M \to \mathbf{R}$, 使得对所有的 $\phi \in T_{\boldsymbol{z}}M$ 都成立 $\Omega(X(\boldsymbol{z}), \phi) = \delta\mathcal{H}(\boldsymbol{z}) \cdot \phi$, 则称 X 是一个 **Hamilton 向量场**. 在这种情况下, 我们用 $X_{\mathcal{H}}$ 来表示 Hamilton 向量场.

假如 ϕ_t 是一个局部 Hamilton 向量场 $X_{\mathcal{H}}$ 的无穷维相流, ϕ_t^* 是它的拉回映射, 则该相流满足 $\phi_t^*\Omega = \Omega$. 对应向量场 $X_{\mathcal{H}}$ 的方程

$$\frac{\partial \phi_t(\boldsymbol{z})}{\partial t} = X_{\mathcal{H}}(\phi_t(\boldsymbol{z}))$$

就是经典 Hamilton 系统, 该系统也可以通过局部坐标表示如下:

$$\frac{\partial \boldsymbol{q}}{\partial t} = \frac{\delta \mathcal{H}}{\delta \boldsymbol{p}}, \quad \frac{\partial \boldsymbol{p}}{\partial t} = -\frac{\delta \mathcal{H}}{\delta \boldsymbol{q}}. \tag{7.0.1}$$

下面为了简单起见, 我们用 x 表示空间变量. 经典 Hamilton 系统 (7.0.1) 保持 Hamilton 辛结构, 即保持

$$\Omega = \int_x \delta \boldsymbol{p} \wedge \delta \boldsymbol{q}, \tag{7.0.2}$$

此处 δ 在无限维情形下表示变分微分算子, 在有限维情形下表示普通微分算子. 向量场 $X_{\mathcal{H}} = \left(\frac{\delta \mathcal{H}}{\delta \boldsymbol{p}}, -\frac{\delta \mathcal{H}}{\delta \boldsymbol{q}}\right)$ 的相流 ϕ_t 是个 Hamilton 辛映射, 即 $\Omega(X_{\mathcal{H}}, \phi) = \delta\mathcal{H} \cdot \phi$. 在局部坐标下, 该性质也可以表示为

$$\left(\frac{\partial \phi_t(\boldsymbol{z})}{\partial \boldsymbol{z}}\right)^{\mathrm{T}} \boldsymbol{J} \frac{\partial \phi_t(\boldsymbol{z})}{\partial \boldsymbol{z}} = \boldsymbol{J}, \tag{7.0.3}$$

其中

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I}_n \\ -\boldsymbol{I}_n & \boldsymbol{0} \end{pmatrix}.$$

经典 Hamilton 系统 (7.0.1) 的保结构算法可以分为两大类: 第一类是, 首先把方程沿着空间方向离散得到一个半离散系统, 加上合适的边界条件使之成为一个有限维 Hamilton 系统, 即 Hamilton 常微分方程组, 然后用 Hamilton 辛算法, 也就是第二章介绍的 Hamilton 辛差分算法, 去数值求解这个方程组. 这类算法只有在适当的边界条件下, 才会沿时间方向保持 Hamilton 辛结构. 第二类是所谓的 Hamilton 多辛算法, 比如多辛 Runge-Kutta 方法. 这类算法是基于方程 (组) 的多辛 Hamilton 形式的, 这种方程 (组) 形式保持一个 Hamilton 多辛结构, 简单地讲, 就是沿时间和空间各有一个辛结构. 这两类算法保持的都是有限维的结构, 前者是在离散的有限维空间上保持辛结构, 而后者是在有限维的构形空间丛上保持多辛结构, 因此这两类算法都不保持无限维的 Hamilton 辛结构 (7.0.2). 和其他非辛算法相比, 这两类算法在长时间计算时的优势是非常明显的. 关于这两种算法有大量的研究成果, 有兴趣的读者可参考原始论文 [Tang93, Shang94, McLachlan94, Bri97, MPS98, MS99, BR01, Reich02, HLW02, Olver04, Cano06, HS08] 等.

在下面展开的内容中, 我们首先定义无限维 Birkhoff 系统, 证明它的解保持无限维 Birkhoff 辛结构; 然后给出一般 Birkhoff 辛结构和四类生成泛函的关系, 以及与生成泛函对应的广义 Hamilton-Jacobi 方程, 从而建立生成泛函方法的理论框架. 在此理论基础上, 以声波方程和电磁场方程为例, 给出一个利用生成泛函构造保无限维 Birkhoff 辛结构算法的系统方法.

§7.1 无限维 Birkhoff 系统

我们以一个波动方程为例来说明什么是无限维 Birkhoff 系统. 考虑二维均匀媒质中带耗散项的声波方程

$$u_{tt} - c(u_{xx} + u_{yy}) + u + \lambda u_t = 0, \tag{7.1.1}$$

其中 $\sqrt{c}$ 是波速, λ 是耗散常数, u_t 表示外力作用耗散项. 我们引入函数 p, 使得 $u_t = p$, 并取一个泛函

$$\mathcal{B}(u,p) = \frac{1}{2}\mathrm{e}^{\lambda t}\iint_{(x,y)}\left(u^2 + c(u_x^2 + u_y^2) + p^2 + \lambda up\right)\mathrm{d}x\mathrm{d}y. \tag{7.1.2}$$

如果令 $\boldsymbol{z} = (u,p)^{\mathrm{T}}$, 通过泛函 (7.1.2), 方程 (7.1.1) 可以表示为下面的形式:

$$\boldsymbol{K}(\boldsymbol{z},x,y,t)\frac{\partial \boldsymbol{z}}{\partial t} = \frac{\delta\mathcal{B}(\boldsymbol{z},x,y,t)}{\delta \boldsymbol{z}} + \frac{\partial \boldsymbol{F}(\boldsymbol{z},x,y,t)}{\partial t}, \tag{7.1.3}$$

其中

$$\boldsymbol{K}(\boldsymbol{z},x,y,t) = \begin{pmatrix} 0 & -\mathrm{e}^{\lambda t} \\ \mathrm{e}^{\lambda t} & 0 \end{pmatrix}, \tag{7.1.4}$$

向量函数 $\boldsymbol{F}(\boldsymbol{z},x,y,t) = \left(\frac{1}{2}\mathrm{e}^{\lambda t}p, -\frac{1}{2}\mathrm{e}^{\lambda t}u\right)^{\mathrm{T}}$, 而 $\frac{\delta\mathcal{B}}{\delta\boldsymbol{z}} = \left(\frac{\delta\mathcal{B}}{\delta u}, \frac{\delta\mathcal{B}}{\delta p}\right)^{\mathrm{T}}$ 表示泛函 $\mathcal{B}(\boldsymbol{z},x,y,t)$ 的变分梯度函数.

我们称具有形式 (7.1.3) 的方程组为**无限维 Birkhoff 系统**. 无限维 Birkhoff 系统可以看做有限维 Birkhoff 系统从有限维空间到无限维空间的一个推广, 这也是我们为何也称方程组 (7.1.3) 为 Birkhoff 系统的原因: 尽管经典 Birkhoff 力学中并无无限维一说, 但我们将方程组 (7.1.3) 视做有限维 Birkhoff 系统的延续, 因此也以 Birkhoff 命名. 无限维 Birkhoff 系统也可以看做无限维 Hamilton 系统从保守系统到非保守系统的一个推广. 为什么这么说呢? 如果我们令 $\lambda = 0$, 那么方程 (7.1.1) 就是没有耗散项的一个保守系统, 它对应的形式 (7.1.3) 就是一个无限维 Hamilton 系统.

首先介绍下面要用到的记号: V 表示 Euclid 空间 $\mathbf{R}^m$ 的一子空间, 局部坐标为 $\boldsymbol{x} = (x_1, \cdots, x_m)$, 代表自变量空间; P 是 $\mathbf{R}^{2n}$ 的开子集, 局

部坐标为 $\boldsymbol{z}=(z_1,\cdots,z_{2n})$, 代表因变量空间. 记偏导数 $\dfrac{\partial^J \boldsymbol{z}}{\partial x_1^{j_1}\cdots\partial x_m^{j_m}}$ 为 $\boldsymbol{z}^{\boldsymbol{j}}$, $\boldsymbol{j}=(j_1,\cdots,j_m)$ 是一个 m 维有序非负整数组, 即 $\boldsymbol{z}^{\boldsymbol{j}}$ 表示 $\boldsymbol{z}$ 分别关于 x_k 求 j_k 阶偏导数, $J=\sum\limits_{k=1}^{m} j_k$ 表示 $\boldsymbol{z}^{\boldsymbol{j}}$ 的偏导数总阶数. 类似地, 用 $z_\mu^{\boldsymbol{j}}$ 表示 $\dfrac{\partial^J z_\mu}{\partial x_1^{j_1}\cdots\partial x_m^{j_m}}(\mu=1,\cdots,2n)$, 用 $\mathrm{D}^{\boldsymbol{j}}=\dfrac{\partial^J}{\partial x_1^{j_1}\cdots\partial x_m^{j_m}}$ 表示 J 阶全微分算子. 记 $\boldsymbol{z}^0=(\boldsymbol{x},\boldsymbol{z})$. 用 $J^\alpha P$ 表示基础空间 $V\times P$ 的 α 阶 Jet 空间, $(\boldsymbol{x},\boldsymbol{z}^{\boldsymbol{j}})_{0\leqslant J\leqslant\alpha}$ 为它的坐标, 并用 $J_{\boldsymbol{z}^0}P$ 和 JP 分别表示 $V\times P$ 的任意有限阶 Jet 空间和 Jet 丛. 下面给出变分微分 1-形式的定义和变分外微分算子 δ 的定义.

定义 7.1.1　如果 $\mathcal{B}=\displaystyle\int_{\boldsymbol{x}} B(\boldsymbol{x},\boldsymbol{z}^{\boldsymbol{j}})\mathrm{d}\boldsymbol{x}$ 是定义在 JP 上的一个泛函, 那么对 $\mathcal{B}$ 作变分, 得到

$$(\delta\mathcal{B})(\delta\boldsymbol{z})=\int_{\boldsymbol{x}}\frac{\partial B}{\partial z_\mu^{\boldsymbol{j}}}\mathrm{d}z_\mu^{\boldsymbol{j}}\mathrm{d}\boldsymbol{x}=\int_{\boldsymbol{x}}\frac{\delta\mathcal{B}}{\delta\boldsymbol{z}}\delta\boldsymbol{z}+\int_{\partial\boldsymbol{x}}L\cdot\mathrm{d}S, \tag{7.1.5}$$

其中 L 是边界项. 显然有 $\dfrac{\delta\mathcal{B}}{\delta\boldsymbol{z}}=\sum\limits_{J\geqslant0}(-1)^J\mathrm{D}^{\boldsymbol{j}}\dfrac{\partial B}{\partial\boldsymbol{z}^{\boldsymbol{j}}}$. 把 $\dfrac{\delta\mathcal{B}}{\delta\boldsymbol{z}}$ 称为泛函 $\mathcal{B}$ 的**变分梯度函数**.

在上面的定义中, $\sum\limits_{J\geqslant0}(-1)^J\mathrm{D}^{\boldsymbol{j}}\dfrac{\partial}{\partial\boldsymbol{z}^{\boldsymbol{j}}}$ 就是 Euler 微分算子, $\dfrac{\partial B}{\partial z_\mu^i}\mathrm{d}z_\mu^i$ 是定义在 Jet 丛 JP 上的延拓 1-形式.

定义 7.1.2　假设 $f_\mu^{\boldsymbol{\nu}}(\boldsymbol{x},\boldsymbol{z}^{\boldsymbol{j}})$ ($\mu=1,\cdots,2n,\boldsymbol{\nu}$ 是任意 m 维的有序非负整数组) 是定义在 Jet 丛 JP 上的函数, $\Theta=\displaystyle\int_{\boldsymbol{x}} f_\mu^{\boldsymbol{\nu}}\mathrm{d}z_\mu^{\boldsymbol{\nu}}\mathrm{d}\boldsymbol{x}$ 是一个变分微分 1-形式. Θ 的**变分外微分**定义如下:

$$\delta\Theta=\int_{\boldsymbol{x}}\mathrm{d}f_\mu^{\boldsymbol{\nu}}\wedge\mathrm{d}z_\mu^{\boldsymbol{\nu}}\mathrm{d}\boldsymbol{x}, \tag{7.1.6}$$

其中 $\mathrm{d}f_\mu^{\boldsymbol{\nu}}=\dfrac{\partial f_\mu^{\boldsymbol{\nu}}}{\partial\boldsymbol{z}^{\boldsymbol{\tau}}}\mathrm{d}\boldsymbol{z}^{\boldsymbol{\tau}}$.

定理 7.1.1　变分 $\delta\mathcal{B}$ 的变分外微分恒为零, 即 $\delta^2\mathcal{B}=0$.

利用 $\dfrac{\partial^2 B}{\partial \boldsymbol{z}^\tau \partial \boldsymbol{z}^\nu}$ 的对称性可以很直接地证明本定理, 因此这里定理证明从略.

我们将采用以下顺序给出和 Birkhoff 系统相关的基本概念: 首先, 介绍 Birkhoff 辛形式和 Birkhoff 辛映射; 然后, 给出 Birkhoff 泛函和辛向量场的定义以及它们之间的关系; 最后, 定义 Birkhoff 系统, 并给出 Birkhoff 辛守恒律.

假设构形空间 $P = \mathcal{F}(\mathbf{R}^3 \times \mathbf{R})$ 是 $\mathbf{R}^{2n}$ 的一个子空间, 它的局部坐标为 $\boldsymbol{z} = (z_1, \cdots, z_{2n})$. $\boldsymbol{K}(\boldsymbol{z}, \boldsymbol{x}, t)$ 表示任意非退化矩阵, 它决定一个非退化变分 2-形式 $\Omega_{\boldsymbol{K}}$:

$$\Omega_{\boldsymbol{K}} = \int_{\boldsymbol{x}} \frac{1}{2}(\boldsymbol{K}(\boldsymbol{z}, \boldsymbol{x}, t)\delta \boldsymbol{z})^{\mathrm{T}} \wedge \delta \boldsymbol{z} = \int_{\boldsymbol{x}} \frac{1}{2}\delta \boldsymbol{z}^{\mathrm{T}} \wedge (\boldsymbol{K}(\boldsymbol{z}, \boldsymbol{x}, t)^{\mathrm{T}}\delta \boldsymbol{z}). \quad (7.1.7)$$

对 P 上的任意向量场 v 和 w, 有

$$\Omega_{\boldsymbol{K}}(v, w) = \int_{\boldsymbol{x}} \frac{1}{2}(\delta \boldsymbol{z}^{\mathrm{T}} \wedge (\boldsymbol{K}(\boldsymbol{z}, \boldsymbol{x}, t)^{\mathrm{T}}\delta \boldsymbol{z}) \cdot (v, w))\mathrm{d}\boldsymbol{x}.$$

定义 7.1.3 假设 $\Omega_{\boldsymbol{K}}$ 是 P 上的一个非退化变分 2-形式. 如果 $\Omega_{\boldsymbol{K}}$ 在 P 上是闭的, 即 $\delta\Omega_{\boldsymbol{K}} = 0$, 则称 $\Omega_{\boldsymbol{K}}$ 是 **Birkhoff 辛形式**或**辛结构**, 也称为 **$\boldsymbol{K}(\boldsymbol{z}, \boldsymbol{x}, t)$-辛形式**. 称带有 Birkhoff 辛结构的空间 $(P, \Omega_{\boldsymbol{K}})$ 为 **Birkhoff 辛空间**或 **$\boldsymbol{K}(\boldsymbol{z}, \boldsymbol{x}, t)$-辛空间**.

因为 $\Omega_{\boldsymbol{K}}$ 的变分外微分为

$$\delta\Omega_{\boldsymbol{K}} = \frac{1}{2}\int_{\boldsymbol{x}} \frac{\partial K_{\mu\nu}}{\partial z_\tau}\delta z_\mu \wedge \delta z_\nu \wedge \delta z_\tau, \quad (7.1.8)$$

所以变分 2-形式 $\Omega_{\boldsymbol{K}}$ 的非退化性和辛性质局部等价于有限维微分 2-形式 $\omega_{\boldsymbol{K}} = \dfrac{1}{2}\delta \boldsymbol{z}^{\mathrm{T}} \wedge (\boldsymbol{K}^{\mathrm{T}}\delta \boldsymbol{z})$ 的非退化性和辛性质. 而非退化闭 2-形式 $\omega_{\boldsymbol{K}}$ 意味着对应矩阵 $\boldsymbol{K}$ 是非退化的, 且满足条件

$$K_{\mu\nu} + K_{\nu\mu} = 0, \quad \frac{\partial K_{\mu\nu}}{\partial z_\tau} + \frac{\partial K_{\nu\tau}}{\partial z_\mu} + \frac{\partial K_{\tau\mu}}{\partial z_\nu} = 0,$$
$$\mu, \nu, \tau = 1, \cdots, 2n. \quad (7.1.9)$$

关于有限维 Hamilton 系统和 Birkhoff 系统的非退化闭 2-形式, 读者可参考文献 [MS99, San83b]. 有了 Birkhoff 辛形式和辛空间的概念, 我们便可定义 Birkhoff 辛映射.

定义 7.1.4 假设 ϕ_t 是一个从 Birkhoff 辛空间 $(P, \Omega_{\boldsymbol{K}})$ 到它自身的映射. 如果对任意 $0 < t \in \mathbf{R}$, ϕ_t 保持 Birkhoff 辛结构 $\Omega_{\boldsymbol{K}}$, 即

$$\phi_t^*\left(\int_{\boldsymbol{x}} \omega_{\boldsymbol{K}}\right) = \int_{\boldsymbol{x}} \omega_{\boldsymbol{K}}, \tag{7.1.10}$$

则称 ϕ_t 是 **Birkhoff 辛映射**或 $\boldsymbol{K}(\boldsymbol{z}, \boldsymbol{x}, t)$**-辛映射**.

定义 7.1.5 假设 $(P, \Omega_{\boldsymbol{K}})$ 是一个 Birkhoff 辛空间, 并且

$$\int_{\boldsymbol{x}} \delta(\boldsymbol{F}(\boldsymbol{z}, \boldsymbol{x}, t)\delta \boldsymbol{z}) = \Omega_{\boldsymbol{K}};$$

X 是 P 上的一个向量场. 令

$$\boldsymbol{Q}(\boldsymbol{z}, \boldsymbol{x}, t) = X(\boldsymbol{z}, \boldsymbol{x}, t) - \boldsymbol{K}(\boldsymbol{z}, \boldsymbol{x}, t)^{-1}\frac{\partial \boldsymbol{F}(\boldsymbol{z}, \boldsymbol{x}, t)}{\partial t}.$$

如果对任意 $\boldsymbol{z} \in P$, 存在泛函 $\mathcal{B}: P \to \mathbf{R}$, 使得

$$\Omega_{\boldsymbol{K}}(\boldsymbol{Q}(\boldsymbol{z}, \boldsymbol{x}, t), \delta \boldsymbol{v}) = \delta\mathcal{B}(\boldsymbol{z}, \boldsymbol{x}, t)(\delta \boldsymbol{v}), \quad \forall \delta \boldsymbol{v} \in P, \tag{7.1.11}$$

则称 X 为 **Birkhoff 辛向量场**或 $\boldsymbol{K}(\boldsymbol{z}, \boldsymbol{x}, t)$**-辛向量场**, 记为 $X_{\mathcal{B}}$, 并称 $\mathcal{B}(\boldsymbol{z}, \boldsymbol{x}, t)$ 为 X 的一个 **Birkhoff 泛函**.

因为矩阵 $\boldsymbol{K}(\boldsymbol{z}, \boldsymbol{x}, t)$ 非退化且满足条件 (7.1.9), 所以存在向量函数 $\boldsymbol{F}(\boldsymbol{z}, \boldsymbol{x}, t)$, 使得

$$\boldsymbol{K}(\boldsymbol{z}, \boldsymbol{x}, t) = \left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{z}}\right)^{\mathrm{T}} - \left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{z}}\right).$$

很显然, $\boldsymbol{F}(\boldsymbol{z}, \boldsymbol{x}, t)$ 不唯一. 如果令

$$\mathcal{B}(\boldsymbol{z}, \boldsymbol{x}, t) = \int_{\boldsymbol{x}} B(\boldsymbol{z}, \boldsymbol{x}, t)\mathrm{d}\boldsymbol{x},$$

那么函数 $B(\boldsymbol{z}, \boldsymbol{x}, t)$ 也不唯一. 我们采用经典同伦运算得到 $\boldsymbol{F}(\boldsymbol{z}, \boldsymbol{x}, t)$ 和 $\mathcal{B}(\boldsymbol{z}, \boldsymbol{x}, t)$:

$$F_\tau(\boldsymbol{z}, \boldsymbol{x}, t) = \int_0^1 \lambda z_\nu K_{\nu\tau}(\lambda \boldsymbol{z}, \boldsymbol{x}, t)\mathrm{d}\lambda,$$

$$\mathcal{B}(\boldsymbol{z}, \boldsymbol{x}, t) = \int_0^1 \boldsymbol{z}^{\mathrm{T}}(\boldsymbol{K}\boldsymbol{Q})(\lambda \boldsymbol{z}, \boldsymbol{x}, t)\mathrm{d}\lambda.$$

命题 7.1.1 已知一个泛函 $\mathcal{B}(\boldsymbol{z},\boldsymbol{x},t)$, 它的 $\boldsymbol{K}(\boldsymbol{z},\boldsymbol{x},t)$-辛向量场 $X_{\mathcal{B}}$ 如下所示:

$$X_{\mathcal{B}} = \boldsymbol{K}^{-1}\frac{\delta\mathcal{B}}{\delta\boldsymbol{z}} + \boldsymbol{K}^{-1}\frac{\partial \boldsymbol{F}(\boldsymbol{z},\boldsymbol{x},t)}{\partial t}. \tag{7.1.12}$$

证明 对任意 $\delta\boldsymbol{v}\in P$, 有

$$\begin{aligned}
&\Omega_{\boldsymbol{K}}\left(X_{\mathcal{B}} - \boldsymbol{K}^{-1}\frac{\partial \boldsymbol{F}(\boldsymbol{z},\boldsymbol{x},t)}{\partial t}, \delta\boldsymbol{v}\right)\\
&= \frac{1}{2}\int_{\boldsymbol{x}}(\boldsymbol{K}\delta\boldsymbol{z})^{\mathrm{T}}\wedge\delta\boldsymbol{z}\left(\boldsymbol{K}^{-1}\frac{\delta\mathcal{B}}{\delta\boldsymbol{z}}, \delta\boldsymbol{v}\right)\mathrm{d}\boldsymbol{x}\\
&= \frac{1}{2}\int_{\boldsymbol{x}}\left(\left(\boldsymbol{K}\boldsymbol{K}^{-1}\frac{\delta\mathcal{B}}{\delta\boldsymbol{z}}\right)^{\mathrm{T}}\delta\boldsymbol{v} - (\boldsymbol{K}\delta\boldsymbol{v})^{\mathrm{T}}\boldsymbol{K}^{-1}\frac{\delta\mathcal{B}}{\delta\boldsymbol{z}}\right)\mathrm{d}\boldsymbol{x}\\
&= \frac{1}{2}\int_{\boldsymbol{x}}\left((\delta\boldsymbol{v})^{\mathrm{T}}\left(\frac{\delta\mathcal{B}}{\delta\boldsymbol{z}}\right) - (\delta\boldsymbol{v})^{\mathrm{T}}\boldsymbol{K}^{\mathrm{T}}\boldsymbol{K}^{-1}\frac{\delta\mathcal{B}}{\delta\boldsymbol{z}}\right)\mathrm{d}\boldsymbol{x}\\
&= \int_{\boldsymbol{x}}(\delta\boldsymbol{v})^{\mathrm{T}}\frac{\delta\mathcal{B}}{\delta\boldsymbol{z}}\mathrm{d}\boldsymbol{x} = \delta\mathcal{B}(\boldsymbol{z},\boldsymbol{x},t)(\delta\boldsymbol{v}).
\end{aligned}$$

根据定义 7.1.5 知道, (7.1.12) 式是 $\mathcal{B}$ 的 $\boldsymbol{K}(\boldsymbol{z},\boldsymbol{x},t)$-辛向量场. □

定义 7.1.6 假设 $\boldsymbol{K}(\boldsymbol{z},\boldsymbol{x},t)$ 是满足条件 (7.1.9) 的非退化矩阵, $X_{\mathcal{B}}$ 是泛函 $\mathcal{B}(\boldsymbol{z},\boldsymbol{x},t)$ 的 $\boldsymbol{K}(\boldsymbol{z},\boldsymbol{x},t)$-辛向量场, 则称方程

$$\frac{\partial \boldsymbol{z}}{\partial t} = X_{\mathcal{B}}(\boldsymbol{z},\boldsymbol{x},t) \tag{7.1.13}$$

为**无限维 Birkhoff 系统**.

定理 7.1.2 无限维 Birkhoff 系统 (7.1.3) 的解是 Birkhoff 辛映射.

证明 为了简单起见, 我们把 $\dfrac{\partial \boldsymbol{z}}{\partial t}$ 记做 $\dot{\boldsymbol{z}}$. 方程 (7.1.13) 两端同时乘以 $\delta\boldsymbol{z}$, 且关于空间变量 $\boldsymbol{x}$ 积分, 得到

$$\int_{\boldsymbol{x}}\boldsymbol{K}\dot{\boldsymbol{z}}\delta\boldsymbol{z} = \int_{\boldsymbol{x}}X_{\mathcal{B}}\delta\boldsymbol{z} = \int_{\boldsymbol{x}}\frac{\delta\mathcal{B}}{\delta\boldsymbol{z}}\delta\boldsymbol{z} + \int_{\boldsymbol{x}}\frac{\partial\boldsymbol{F}}{\partial t}\delta\boldsymbol{z}. \tag{7.1.14}$$

显然有

$$\int_{\boldsymbol{x}}\frac{\delta\mathcal{B}}{\delta\boldsymbol{z}}\delta\boldsymbol{z} = \int_{\boldsymbol{x}}\frac{\partial B}{\partial\boldsymbol{z}}\mathrm{d}\boldsymbol{z}\mathrm{d}\boldsymbol{x} + \int_{\partial\boldsymbol{x}}L\cdot\mathrm{d}S,$$

其中 L 为分部积分所得的边界项. 我们不妨设边界条件满足

$$\int_{\partial\boldsymbol{x}}L\cdot\mathrm{d}S = 0,$$

那么

$$\int_{\boldsymbol{x}} \boldsymbol{K}\dot{\boldsymbol{z}}\delta\boldsymbol{z} = \int_{\boldsymbol{x}} \frac{\partial \mathcal{B}}{\partial \boldsymbol{z}} \mathrm{d}\boldsymbol{z}\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{x}} \frac{\partial \boldsymbol{F}}{\partial t}\delta\boldsymbol{z}. \tag{7.1.15}$$

对 (7.1.15) 式两端作变分, 得到

$$\int_{\boldsymbol{x}} \left(\delta\boldsymbol{z} \wedge \frac{\partial \boldsymbol{K}}{\partial \boldsymbol{z}}\dot{\boldsymbol{z}}\delta\boldsymbol{z} + \delta\dot{\boldsymbol{z}} \wedge \boldsymbol{K}\delta\boldsymbol{z}\right) = \delta^2\mathcal{B} + \int_{\boldsymbol{x}} \delta\boldsymbol{z} \wedge \frac{\partial \boldsymbol{F}_t}{\partial \boldsymbol{z}}\delta\boldsymbol{z}. \tag{7.1.16}$$

已知 $\delta^2\mathcal{B} = 0$ 和 $\boldsymbol{K} = (\nabla\boldsymbol{F})^{\mathrm{T}} - \nabla\boldsymbol{F}$, 且 $\boldsymbol{K}$ 满足条件 (7.1.9), 所以上式可以改写为

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\boldsymbol{x}} (\delta\boldsymbol{z} \wedge \boldsymbol{K}\delta\boldsymbol{z}) = 0. \tag{7.1.17}$$

这表示辛形式 $\Omega_{\boldsymbol{K}}$ 沿 Birkhoff 系统 (7.1.13) 的解不变, 即系统 (7.1.13) 的解是 $\boldsymbol{K}(\boldsymbol{z},\boldsymbol{x},t)$-辛映射. □

注 7.1.1　定理 7.1.2 证明过程中假设的边界条件不影响方程沿时间方向的离散方法的选择, 因此在本章我们假设这样的边界条件始终存在. 这个假设可保证泛函的变分梯度的变分导数是对称的.

§7.2　$\boldsymbol{K}(x,t)$-辛结构和生成泛函

在本节中, 我们将给出只依赖于自变量 (x,t) 的 $1+1$ 维 Brikhoff 系统的 Birkhoff 辛结构和四类生成泛函之间的关系. 下面我们假设 Birkhoff 系统的解 $\hat{\boldsymbol{z}} = (\hat{\boldsymbol{u}}, \hat{\boldsymbol{p}})$ 在初始时刻 t_0 的值为 $\boldsymbol{z} = (\boldsymbol{u}, \boldsymbol{p})$.

假设 $\widehat{\boldsymbol{K}}$ 是只依赖于自变量 (x,t) 的 Birkhoff 辛矩阵, 且 $\widehat{\boldsymbol{K}}$ 可以表示为分块矩阵:

$$\widehat{\boldsymbol{K}}(x,t) = \begin{pmatrix} \boldsymbol{0} & -\widehat{\boldsymbol{A}}(x,t) \\ \widehat{\boldsymbol{A}}(x,t) & \boldsymbol{0} \end{pmatrix}. \tag{7.2.1}$$

因为 $\widehat{\boldsymbol{K}}$ 是反对称的, 所以 $\widehat{\boldsymbol{A}}$ 是对称矩阵, 对应的 Birkhoff 系统则可以表示为

$$\begin{aligned} -\widehat{\boldsymbol{A}}\frac{\partial \hat{\boldsymbol{p}}}{\partial t} &= \frac{\delta \mathcal{B}}{\delta \hat{\boldsymbol{u}}} + \frac{\partial \boldsymbol{F}_1}{\partial t}, \\ \widehat{\boldsymbol{A}}\frac{\partial \hat{\boldsymbol{u}}}{\partial t} &= \frac{\delta \mathcal{B}}{\delta \hat{\boldsymbol{p}}} + \frac{\partial \boldsymbol{F}_2}{\partial t}. \end{aligned} \tag{7.2.2}$$

令向量函数 $\boldsymbol{F}=(\boldsymbol{F}_1^{\mathrm{T}},\boldsymbol{F}_2^{\mathrm{T}})^{\mathrm{T}}$ 和 $\widehat{\boldsymbol{K}}$ 之间满足下面的关系:

$$\widehat{\boldsymbol{A}}=\frac{\partial \boldsymbol{F}_1}{\partial \widehat{\boldsymbol{p}}}-\frac{\partial \boldsymbol{F}_2}{\partial \widehat{\boldsymbol{u}}}. \tag{7.2.3}$$

由 $\widehat{\boldsymbol{K}}$ 可诱导出 2-形式

$$\Omega_{\widehat{\boldsymbol{K}}}=\int_x \delta\widehat{\boldsymbol{u}}\wedge\widehat{\boldsymbol{A}}\delta\widehat{\boldsymbol{p}}. \tag{7.2.4}$$

矩阵 $\widehat{\boldsymbol{A}}(x,t)$ 和 $\widehat{\boldsymbol{K}}(x,t)$ 在初始时刻 t_0 分别为 $\boldsymbol{A}(x,t_0)$ 和 $\boldsymbol{K}(x,t_0)$. 因为 (7.2.1) 式中的 $\widehat{\boldsymbol{K}}(x,t)$ 是满足条件 (7.1.9) 的非退化矩阵, 所以 Birkhoff 系统 (7.2.2) 的解 $\phi_t:(\boldsymbol{u},\boldsymbol{p})\mapsto(\widehat{\boldsymbol{u}},\widehat{\boldsymbol{p}})$ 是一个 $\widehat{\boldsymbol{K}}(x,t)$-辛映射.

注 7.2.1 由于矩阵 $\widehat{\boldsymbol{K}}$ 满足条件 (7.1.9), 我们可以通过令

$$\widehat{\boldsymbol{A}}(x,t)=\frac{\partial \boldsymbol{F}_1}{\partial \widehat{\boldsymbol{p}}}-\frac{\partial \boldsymbol{F}_2}{\partial \widehat{\boldsymbol{u}}}$$

来构造 $\widehat{\boldsymbol{K}}$. 向量函数 $\boldsymbol{F}$ 不唯一, 我们不妨这样选择 $\boldsymbol{F}$: $\boldsymbol{F}_1$ 不显含 $\widehat{\boldsymbol{u}}$, 但关于 $\widehat{\boldsymbol{p}}$ 是线性函数; $\boldsymbol{F}_2$ 不显含 $\widehat{\boldsymbol{p}}$, 但关于 $\widehat{\boldsymbol{u}}$ 是线性函数.

7.2.1 $\widehat{\boldsymbol{K}}$-辛结构和生成泛函 $\mathcal{S}$ 的关系

为了简便, 我们用 $\mathcal{S}(\boldsymbol{u},\widehat{\boldsymbol{u}},t)$ 表示 $\mathcal{S}(\boldsymbol{u},\widehat{\boldsymbol{u}},\boldsymbol{u}_x,\widehat{\boldsymbol{u}}_x,\boldsymbol{u}_{xx},\widehat{\boldsymbol{u}}_{xx},\cdots,t)$, 用 $\mathcal{S}^2(\widehat{\boldsymbol{u}},\boldsymbol{p},t)$, $\mathcal{S}^1(\boldsymbol{u},\widehat{\boldsymbol{p}},t)$, $\mathcal{S}^3(\boldsymbol{v},\boldsymbol{w},t)$ 分别表示

$$\mathcal{S}^1(\boldsymbol{u},\widehat{\boldsymbol{p}},\boldsymbol{u}_x,\widehat{\boldsymbol{p}}_x,\boldsymbol{u}_{xx},\widehat{\boldsymbol{p}}_{xx},\cdots,t),\quad \mathcal{S}^2(\widehat{\boldsymbol{u}},\boldsymbol{p},\widehat{\boldsymbol{u}}_x,\boldsymbol{p}_x,\widehat{\boldsymbol{u}}_{xx},\boldsymbol{p}_{xx},\cdots,t),$$
$$\mathcal{S}^3(\boldsymbol{v},\boldsymbol{w},\boldsymbol{v}_x,\boldsymbol{w}_x,\boldsymbol{v}_{xx},\boldsymbol{w}_{xx},\cdots,t).$$

由定义 7.1.4 知道, 一个 $\widehat{\boldsymbol{K}}$- 辛映射 ϕ_t 一定保持 $\widehat{\boldsymbol{K}}$-辛结构, 所以有

$$\Omega_{\widehat{\boldsymbol{K}}}=\Omega_{\boldsymbol{K}},\quad \text{即}\quad \int_x \delta\widehat{\boldsymbol{u}}\wedge\widehat{\boldsymbol{A}}\delta\widehat{\boldsymbol{p}}=\int_x \delta\boldsymbol{u}\wedge\boldsymbol{A}\delta\boldsymbol{p}. \tag{7.2.5}$$

此性质等价于

$$\int_x(\delta\widehat{\boldsymbol{u}}\wedge\widehat{\boldsymbol{A}}\delta\widehat{\boldsymbol{p}}-\delta\boldsymbol{u}\wedge\boldsymbol{A}\delta\boldsymbol{p})=0. \tag{7.2.6}$$

这意味着 (7.2.6) 式左端是一个二阶变分, 所以有下面的结论:

定理 7.2.1 给定映射 $\phi_t : (\boldsymbol{u}, \boldsymbol{p}) \mapsto (\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{p}})$, 假设存在一个泛函 $\mathcal{S}(\boldsymbol{u}, \widehat{\boldsymbol{u}}, t)$, 使得

$$\int_x ((\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}})^{\mathrm{T}} \delta\widehat{\boldsymbol{u}} - (\boldsymbol{A}\boldsymbol{p})^{\mathrm{T}} \delta\boldsymbol{u}) = \delta\mathcal{S}(\boldsymbol{u}, \widehat{\boldsymbol{u}}, t), \tag{7.2.7}$$

则 ϕ_t 是一个 $\widehat{\boldsymbol{K}}$-辛映射. 这时称泛函 $\mathcal{S}$ 为辛映射 ϕ_t 的生成泛函.

将 (7.2.7) 式等价变形为

$$\frac{\delta\mathcal{S}}{\delta\widehat{\boldsymbol{u}}} = \widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}}, \quad \frac{\delta\mathcal{S}}{\delta\boldsymbol{u}} = -\boldsymbol{A}\boldsymbol{p}. \tag{7.2.8}$$

对 (7.2.8) 式中的第二个关系式两端关于 t 求偏导数, 得到

$$\frac{\partial}{\partial t}\left(\frac{\delta\mathcal{S}}{\delta\boldsymbol{u}}\right) + \left(\frac{\partial\widehat{\boldsymbol{u}}}{\partial t}\right)^{\mathrm{T}} \frac{\delta^2\mathcal{S}}{\delta\widehat{\boldsymbol{u}}\delta\boldsymbol{u}} = \boldsymbol{0}, \tag{7.2.9}$$

因为注 7.1.1 所假设的边界条件, 所以泛函 $\mathcal{S}$ 的二阶变分导数 $\dfrac{\delta^2\mathcal{S}}{\delta\widehat{\boldsymbol{u}}\delta\boldsymbol{u}}$ 是对称的, 即

$$\frac{\delta^2\mathcal{S}}{\delta\widehat{\boldsymbol{u}}\delta\boldsymbol{u}} = \frac{\delta^2\mathcal{S}}{\delta\boldsymbol{u}\delta\widehat{\boldsymbol{u}}}.$$

将 Birkhoff 方程 (7.2.2) 代入等式 (7.2.9), 得到

$$\frac{\partial}{\partial t}\left(\frac{\delta\mathcal{S}}{\delta\boldsymbol{u}}\right) + \left(\frac{\delta\mathcal{B}}{\delta\widehat{\boldsymbol{p}}} + \frac{\partial\boldsymbol{F}_2}{\partial t}\right)^{\mathrm{T}} \widehat{\boldsymbol{A}}^{-1} \frac{\delta^2\mathcal{S}}{\delta\widehat{\boldsymbol{u}}\delta\boldsymbol{u}} = \boldsymbol{0}, \tag{7.2.10}$$

其中 $\mathcal{B} = \mathcal{B}(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{p}}, t)$, $\boldsymbol{F}_2 = \boldsymbol{F}_2(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{p}}, t)$. 由 (7.2.8) 式中的第一个关系式知道

$$\frac{\delta^2\mathcal{S}}{\delta\boldsymbol{u}\delta\widehat{\boldsymbol{u}}} = \widehat{\boldsymbol{A}}\frac{\delta\widehat{\boldsymbol{p}}}{\delta\boldsymbol{u}}.$$

又因为

$$\frac{\delta^2\mathcal{S}}{\delta\boldsymbol{u}\delta\widehat{\boldsymbol{u}}} = \frac{\delta^2\mathcal{S}}{\delta\widehat{\boldsymbol{u}}\delta\boldsymbol{u}},$$

所以 (7.2.10) 式等价变形为

$$\frac{\partial}{\partial t}\left(\frac{\delta\mathcal{S}}{\delta\boldsymbol{u}}\right) + \frac{\delta\mathcal{B}}{\delta\boldsymbol{u}} + \left(\frac{\partial\boldsymbol{F}_2}{\partial t}\right)^{\mathrm{T}} \frac{\delta\widehat{\boldsymbol{p}}}{\delta\boldsymbol{u}} = \boldsymbol{0}. \tag{7.2.11}$$

现在我们构造如下泛函:

$$\overline{\mathcal{B}} = \int_x \widehat{\boldsymbol{p}}^{\mathrm{T}} \frac{\partial \boldsymbol{F}_2}{\partial t} \mathrm{d}x. \tag{7.2.12}$$

显然 $\overline{\mathcal{B}}$ 满足 $\left(\dfrac{\partial \boldsymbol{F}_2}{\partial t}\right)^{\mathrm{T}} \dfrac{\delta \widehat{\boldsymbol{p}}}{\delta \boldsymbol{u}} = \dfrac{\delta \overline{\mathcal{B}}}{\delta \boldsymbol{u}}$, 所以 (7.2.11) 式可以改写为

$$\frac{\partial}{\partial t}\left(\frac{\delta \mathcal{S}}{\delta \boldsymbol{u}}\right) + \frac{\delta(\mathcal{B}+\overline{\mathcal{B}})}{\delta \boldsymbol{u}} = \boldsymbol{0}. \tag{7.2.13}$$

根据上面的推导, 我们得出以下定理:

定理 7.2.2 已知关于泛函 $\mathcal{S}$ 的偏微分方程

$$\frac{\partial \mathcal{S}}{\partial t} + (\mathcal{B}+\overline{\mathcal{B}})\left(\widehat{\boldsymbol{u}}, \widehat{\boldsymbol{A}}^{-1}\frac{\delta \mathcal{S}}{\delta \widehat{\boldsymbol{u}}}, t\right) = 0 \tag{7.2.14}$$

满足初始条件

$$\frac{\delta}{\delta \boldsymbol{u}}\mathcal{S}(\boldsymbol{u},\boldsymbol{u},t_0) + \frac{\delta}{\delta \widehat{\boldsymbol{u}}}\mathcal{S}(\boldsymbol{u},\boldsymbol{u},t_0) = \boldsymbol{0},$$

如果 $\mathcal{S}(\boldsymbol{u},\widehat{\boldsymbol{u}},t)$ 是它的解, 且 $\dfrac{\delta^2 \mathcal{S}}{\delta \boldsymbol{u}\delta \widehat{\boldsymbol{u}}}$ 是可逆的, 则 (7.2.8) 式决定的映射 $\phi_t : (\boldsymbol{u},\boldsymbol{p}) \mapsto (\widehat{\boldsymbol{u}},\widehat{\boldsymbol{p}})$ 是 Birkhoff 系统 (7.2.2) 的解. 这时称方程 (7.2.14) 为生成泛函 $\mathcal{S}$ 的**广义 Hamilton-Jacobi 方程** (简称 **H-J 方程**).

证明 对方程 (7.2.14) 两端关于 $\boldsymbol{u}$ 求变分导数, 得到

$$\frac{\delta}{\delta \boldsymbol{u}}\left(\frac{\partial \mathcal{S}}{\partial t}\right) + \left(\frac{\delta(\mathcal{B}+\overline{\mathcal{B}})}{\delta \widehat{\boldsymbol{p}}}\right)^{\mathrm{T}} \frac{\delta \widehat{\boldsymbol{p}}}{\delta \boldsymbol{u}} = \boldsymbol{0}. \tag{7.2.15}$$

对 (7.2.8) 式中的第二个关系式两端关于 t 求偏导数, 得到

$$\frac{\partial}{\partial t}\left(\frac{\delta \mathcal{S}}{\delta \boldsymbol{u}}\right) + \left(\frac{\partial \widehat{\boldsymbol{u}}}{\partial t}\right)^{\mathrm{T}} \frac{\delta^2 \mathcal{S}}{\delta \widehat{\boldsymbol{u}}\delta \boldsymbol{u}} = \boldsymbol{0}. \tag{7.2.16}$$

因为

$$\frac{\delta}{\delta \boldsymbol{u}}\left(\frac{\partial \mathcal{S}}{\partial t}\right) = \frac{\partial}{\partial t}\left(\frac{\delta \mathcal{S}}{\delta \boldsymbol{u}}\right),$$

所以上式变形为

$$\frac{\delta}{\delta \boldsymbol{u}}\left(\frac{\partial \mathcal{S}}{\partial t}\right) + \left(\frac{\partial \widehat{\boldsymbol{u}}}{\partial t}\right)^{\mathrm{T}} \frac{\delta^2 \mathcal{S}}{\delta \widehat{\boldsymbol{u}}\delta \boldsymbol{u}} = \boldsymbol{0}. \tag{7.2.17}$$

因为 $\dfrac{\delta^2\mathcal{S}}{\delta\boldsymbol{u}\delta\widehat{\boldsymbol{u}}}$ 是可逆的, 且

$$\frac{\delta\widehat{\boldsymbol{p}}}{\delta\boldsymbol{u}}=\widehat{\boldsymbol{A}}^{-1}\frac{\delta^2\mathcal{S}}{\delta\boldsymbol{u}\delta\widehat{\boldsymbol{u}}},$$

所以 (7.2.17) 式减去 (7.2.15) 式可以得到

$$\widehat{\boldsymbol{A}}\frac{\partial\widehat{\boldsymbol{u}}}{\partial t}=\frac{\delta(\mathcal{B}+\overline{\mathcal{B}})}{\delta\widehat{\boldsymbol{p}}}. \tag{7.2.18}$$

这正是 Birkhoff 系统 (7.2.2) 的第二个方程.

再对 (7.2.8) 式中的第一个关系式两端关于 t 求偏导数, 同时对方程 (7.2.14) 两端关于 $\widehat{\boldsymbol{u}}$ 求变分导数, 分别得到

$$\frac{\partial}{\partial t}\left(\frac{\delta\mathcal{S}}{\delta\widehat{\boldsymbol{u}}}\right)+\left(\frac{\partial\widehat{\boldsymbol{u}}}{\partial t}\right)^{\mathrm{T}}\frac{\delta^2\mathcal{S}}{\delta\widehat{\boldsymbol{u}}^2}-\widehat{\boldsymbol{A}}\frac{\partial\widehat{\boldsymbol{p}}}{\partial t}-\frac{\partial\widehat{\boldsymbol{A}}}{\partial t}\widehat{\boldsymbol{p}}=\boldsymbol{0}, \tag{7.2.19}$$

$$\frac{\delta}{\delta\widehat{\boldsymbol{u}}}\left(\frac{\partial\mathcal{S}}{\partial t}\right)+\frac{\delta(\mathcal{B}+\overline{\mathcal{B}})}{\delta\widehat{\boldsymbol{u}}}+\left(\frac{\delta(\mathcal{B}+\overline{\mathcal{B}})}{\delta\widehat{\boldsymbol{p}}}\right)^{\mathrm{T}}\widehat{\boldsymbol{A}}^{-1}\frac{\delta^2\mathcal{S}}{\delta\widehat{\boldsymbol{u}}^2}=\boldsymbol{0}. \tag{7.2.20}$$

(7.2.20) 式减去 (7.2.19) 式, 得到

$$\widehat{\boldsymbol{A}}\frac{\partial\widehat{\boldsymbol{p}}}{\partial t}+\frac{\partial\widehat{\boldsymbol{A}}}{\partial t}\widehat{\boldsymbol{p}}=-\frac{\delta(\mathcal{B}+\overline{\mathcal{B}})}{\delta\widehat{\boldsymbol{u}}}. \tag{7.2.21}$$

由 $\overline{\mathcal{B}}$ 的构造以及 $\widehat{\boldsymbol{A}}$ 和 $\boldsymbol{F}$ 之间的关系知道, 上面的方程等价于 Birkhoff 系统 (7.2.2) 的第一个方程. □

7.2.2 生成泛函 $\mathcal{S}^1$, $\mathcal{S}^2$ 和 $\mathcal{S}^3$

因为

$$\int_x((\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}})^{\mathrm{T}}\delta\widehat{\boldsymbol{u}}+(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{u}})^{\mathrm{T}}\delta\widehat{\boldsymbol{p}})=\delta\int_x(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}})^{\mathrm{T}}\widehat{\boldsymbol{u}}\mathrm{d}x,$$

所以

$$\int_x((\widehat{\boldsymbol{A}}\widehat{\boldsymbol{u}})^{\mathrm{T}}\delta\widehat{\boldsymbol{p}}+(\boldsymbol{A}\boldsymbol{p})^{\mathrm{T}}\delta\boldsymbol{u})=\delta\left(\int_x(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}})^{\mathrm{T}}\widehat{\boldsymbol{u}}\mathrm{d}x-\mathcal{S}(\boldsymbol{u},\widehat{\boldsymbol{u}},t)\right). \tag{7.2.22}$$

假如存在另一个泛函 $\mathcal{S}^1(\boldsymbol{u},\widehat{\boldsymbol{p}},t)$, 使得

$$\mathcal{S}^1(\boldsymbol{u},\widehat{\boldsymbol{p}},t)=\int_x(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}})^{\mathrm{T}}(\widehat{\boldsymbol{u}}-\boldsymbol{u})\mathrm{d}x-\mathcal{S}(\boldsymbol{u},\widehat{\boldsymbol{u}},t), \tag{7.2.23}$$

则 $\mathcal{S}^1$ 必须满足条件

$$\int_x((\widehat{\boldsymbol{A}}\widehat{\boldsymbol{u}})^{\mathrm{T}}\delta\widehat{\boldsymbol{p}}+(\boldsymbol{A}\boldsymbol{p})^{\mathrm{T}}\delta\boldsymbol{u})=\delta\left(\int_x(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}})^{\mathrm{T}}\boldsymbol{u}\mathrm{d}x+\mathcal{S}^1(\boldsymbol{u},\widehat{\boldsymbol{p}},t)\right),\quad(7.2.24)$$

即 $\mathcal{S}^1$ 有如下变分梯度函数:

$$\frac{\delta\mathcal{S}^1}{\delta\widehat{\boldsymbol{p}}}=\widehat{\boldsymbol{A}}(\widehat{\boldsymbol{u}}-\boldsymbol{u}),\quad\frac{\delta\mathcal{S}^1}{\delta\boldsymbol{u}}=-\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}}+\boldsymbol{A}\boldsymbol{p}.\quad(7.2.25)$$

除了 $\mathcal{S}^1(\boldsymbol{u},\widehat{\boldsymbol{p}},t)$, 我们还可以通过

$$\mathcal{S}^2(\widehat{\boldsymbol{u}},\boldsymbol{p},t)\quad 和\quad\mathcal{S}^3((\widehat{\boldsymbol{u}}+\boldsymbol{u})/2,(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}}+\boldsymbol{A}\boldsymbol{p})/2,t)$$

来构造 Birkhoff 辛映射 $\phi_t:(\boldsymbol{u},\boldsymbol{p})\mapsto(\widehat{\boldsymbol{u}},\widehat{\boldsymbol{p}})$. 并且对任意的 $i=1,2,3$, 当 $\mathcal{S}^i=0$ 时, ϕ_t 收敛到恒等映射.

引理 7.2.1 假如 $\phi_t:(\boldsymbol{u},\boldsymbol{p})\mapsto(\widehat{\boldsymbol{u}},\widehat{\boldsymbol{p}})$ 是收敛到恒等映射的光滑映射. 如果存在泛函 $\mathcal{S}^1(\boldsymbol{u},\widehat{\boldsymbol{p}},t)$, 使得

$$\int_x((\widehat{\boldsymbol{A}}\widehat{\boldsymbol{u}})^{\mathrm{T}}\delta\widehat{\boldsymbol{p}}+(\boldsymbol{A}\boldsymbol{p})^{\mathrm{T}}\delta\boldsymbol{u})=\delta\left(\int_x(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}})^{\mathrm{T}}\boldsymbol{u}\mathrm{d}x+\mathcal{S}^1(\boldsymbol{u},\widehat{\boldsymbol{p}},t)\right),\quad(7.2.26)$$

或者存在泛函 $\mathcal{S}^2(\widehat{\boldsymbol{u}},\boldsymbol{p},t)$, 使得

$$\int_x((\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}})^{\mathrm{T}}\delta\widehat{\boldsymbol{u}}+(\boldsymbol{A}\boldsymbol{u})^{\mathrm{T}}\delta\boldsymbol{p})=\delta\left(\int_x(\boldsymbol{A}\boldsymbol{p})^{\mathrm{T}}\widehat{\boldsymbol{u}}\mathrm{d}x-\mathcal{S}^2(\widehat{\boldsymbol{u}},\boldsymbol{p},t)\right),\quad(7.2.27)$$

或者存在泛函 $\mathcal{S}^3(\boldsymbol{v},\boldsymbol{w},t)$, 使得

$$2\delta\mathcal{S}^3=\int_x(\widehat{\boldsymbol{u}}-\boldsymbol{u})^{\mathrm{T}}\delta(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}}+\boldsymbol{A}\boldsymbol{p})-\int_x(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}}-\boldsymbol{A}\boldsymbol{p})^{\mathrm{T}}\delta(\widehat{\boldsymbol{u}}+\boldsymbol{u}),\quad(7.2.28)$$

其中

$$\boldsymbol{v}=(\widehat{\boldsymbol{u}}+\boldsymbol{u})/2,\quad\boldsymbol{w}=(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}}+\boldsymbol{A}\boldsymbol{p})/2.$$

则 ϕ_t 是 $\widehat{\boldsymbol{K}}$-辛映射. 这时称 $\mathcal{S}^i$ $(i=1,2,3)$ 为 ϕ_t 的**生成泛函**.

证明 前面已经推导过关系式 (7.2.26), 所以我们在这里只证明关系式 (7.2.27) 和 (7.2.28).

因为

$$\int_x(\boldsymbol{A}\boldsymbol{p})^{\mathrm{T}}\delta\boldsymbol{u}=\delta\int_x(\boldsymbol{A}\boldsymbol{p})^{\mathrm{T}}\boldsymbol{u}\mathrm{d}x-\int_x(\boldsymbol{A}\boldsymbol{u})^{\mathrm{T}}\delta\boldsymbol{p},$$

所以如果选择的 $\mathcal{S}^2(\widehat{\boldsymbol{u}},\boldsymbol{p},t)$ 满足

$$\int_x (\boldsymbol{A}\boldsymbol{p})^{\mathrm{T}}\widehat{\boldsymbol{u}}\mathrm{d}x - \mathcal{S}^2(\widehat{\boldsymbol{u}},\boldsymbol{p},t) = \int_x (\boldsymbol{A}\boldsymbol{p})^{\mathrm{T}}\boldsymbol{u}\mathrm{d}x + \mathcal{S}(\boldsymbol{u},\widehat{\boldsymbol{u}},t), \tag{7.2.29}$$

则我们能直接从关系式 (7.2.27) 推出 (7.2.7) 式. 此时, 由定理 7.2.1 知道, ϕ_t 是 $\widehat{\boldsymbol{K}}$-辛映射.

我们把 (7.2.7) 式改写为下面的形式:

$$\begin{aligned}&\delta\left(\int_x (\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}} + \boldsymbol{A}\boldsymbol{p})^{\mathrm{T}}(\widehat{\boldsymbol{u}} - \boldsymbol{u})\mathrm{d}x - 2\mathcal{S}\right)\\ &\quad = \int_x (\widehat{\boldsymbol{u}} - \boldsymbol{u})^{\mathrm{T}}\delta(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}} + \boldsymbol{A}\boldsymbol{p}) - \int_x (\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}} - \boldsymbol{A}\boldsymbol{p})^{\mathrm{T}}\delta(\widehat{\boldsymbol{u}} + \boldsymbol{u}),\end{aligned} \tag{7.2.30}$$

那么可以令 $\mathcal{S}^3(\boldsymbol{v},\boldsymbol{w},t)$ 满足

$$2\mathcal{S}^3(\boldsymbol{v},\boldsymbol{w},t) = \int_x (\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}} + \boldsymbol{A}\boldsymbol{p})^{\mathrm{T}}(\widehat{\boldsymbol{u}} - \boldsymbol{u})\mathrm{d}x - 2\mathcal{S}(\boldsymbol{u},\widehat{\boldsymbol{u}},t), \tag{7.2.31}$$

则当 (7.2.28) 式成立时, 有 (7.2.7) 式成立. 由定理 7.2.1 知道, 此时 ϕ_t 是 $\widehat{\boldsymbol{K}}$-辛映射. □

7.2.3　生成泛函 $\mathcal{S}^1, \mathcal{S}^2$ 和 $\mathcal{S}^3$ 的 H-J 方程

我们对 (7.2.23) 式两端关于 t 求偏导数, 得到

$$\begin{aligned}&\frac{\partial \mathcal{S}^1}{\partial t} + \int_x \left(\frac{\partial \widehat{\boldsymbol{p}}}{\partial t}\right)^{\mathrm{T}} \frac{\delta \mathcal{S}^1}{\delta \widehat{\boldsymbol{p}}}\mathrm{d}x\\ &\quad = \int_x \left(\left(\frac{\partial(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}})}{\partial t}\right)^{\mathrm{T}}(\widehat{\boldsymbol{u}} - \boldsymbol{u}) + (\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}})^{\mathrm{T}}\frac{\partial \widehat{\boldsymbol{u}}}{\partial t}\right)\mathrm{d}x\\ &\qquad - \frac{\partial \mathcal{S}}{\partial t} - \int_x \left(\frac{\partial \widehat{\boldsymbol{u}}}{\partial t}\right)^{\mathrm{T}} \frac{\delta \mathcal{S}}{\delta \widehat{\boldsymbol{u}}}\mathrm{d}x.\end{aligned}$$

将 $\mathcal{S}^1(\boldsymbol{u},\widehat{\boldsymbol{p}},t)$ 的变分梯度函数 (7.2.25) 代入上式, 得到

$$\frac{\partial \mathcal{S}^1}{\partial t} = \int_x \left(\frac{\partial \widehat{\boldsymbol{A}}}{\partial t}\widehat{\boldsymbol{p}}\right)^{\mathrm{T}}(\widehat{\boldsymbol{u}} - \boldsymbol{u})\mathrm{d}x - \frac{\partial \mathcal{S}}{\partial t}. \tag{7.2.32}$$

(7.2.25) 式中的第一个关系式也可以改写为

$$\widehat{\boldsymbol{u}} = \boldsymbol{u} + \widehat{\boldsymbol{A}}^{-1}\frac{\delta \mathcal{S}^1}{\delta \widehat{\boldsymbol{p}}}. \tag{7.2.33}$$

将 (7.2.32) 式和 (7.2.33) 式代入 $\mathcal{S}(\boldsymbol{u},\widehat{\boldsymbol{u}},t)$ 的 H-J 方程 (7.2.14), 则能得到 $\mathcal{S}^1(\boldsymbol{u},\widehat{\boldsymbol{p}},t)$ 的 H-J 方程, 所以我们给出下面的定理.

定理 7.2.3 *假设生成泛函 $\mathcal{S}^1(\boldsymbol{u},\widehat{\boldsymbol{p}},t)$ 是满足初始条件 $\mathcal{S}^1(\boldsymbol{u},\widehat{\boldsymbol{p}},t_0)=0$ 的 H-J 方程*

$$\frac{\partial\mathcal{S}^1}{\partial t}=\int_x\left(\frac{\partial\widehat{\boldsymbol{A}}}{\partial t}\widehat{\boldsymbol{p}}\right)^{\mathrm{T}}\widehat{\boldsymbol{A}}^{-1}\frac{\delta\mathcal{S}^1}{\delta\widehat{\boldsymbol{p}}}\mathrm{d}x+(\mathcal{B}+\overline{\mathcal{B}})\left(\boldsymbol{u}+\widehat{\boldsymbol{A}}^{-1}\frac{\delta\mathcal{S}^1}{\delta\widehat{\boldsymbol{p}}},\widehat{\boldsymbol{p}},t\right)\tag{7.2.34}$$

的解. 如果 $\dfrac{\delta\widehat{\boldsymbol{u}}}{\delta\boldsymbol{u}}$ 是可逆的, 那么 (7.2.25) 式决定的映射 $\phi_t:(\boldsymbol{u},\boldsymbol{p})\mapsto(\widehat{\boldsymbol{u}},\widehat{\boldsymbol{p}})$ 是 Birkhoff 系统 (7.2.2) 的解.

证明 为了简便, 令 $\mathcal{H}=(\mathcal{B}+\overline{\mathcal{B}})$. 对 (7.2.25) 式中的第二个关系式两端关于 t 求偏导数, 再变形整理, 得到

$$\begin{aligned}-\frac{\partial}{\partial t}\left(\frac{\delta\mathcal{S}^1}{\delta\boldsymbol{u}}\right)&=\widehat{\boldsymbol{A}}\frac{\partial\widehat{\boldsymbol{p}}}{\partial t}+\frac{\partial\widehat{\boldsymbol{A}}}{\partial t}\widehat{\boldsymbol{p}}+\frac{\delta^2\mathcal{S}^1}{\delta\widehat{\boldsymbol{p}}\delta\boldsymbol{u}}\frac{\partial\widehat{\boldsymbol{p}}}{\partial t}\\&=\left(\frac{\delta^2\mathcal{S}^1}{\delta\widehat{\boldsymbol{p}}\delta\boldsymbol{u}}\widehat{\boldsymbol{A}}^{-1}+\boldsymbol{I}\right)\widehat{\boldsymbol{A}}\frac{\partial\widehat{\boldsymbol{p}}}{\partial t}+\frac{\partial\widehat{\boldsymbol{A}}}{\partial t}\widehat{\boldsymbol{p}}.\end{aligned}\tag{7.2.35}$$

再对方程 (7.2.34) 两端关于 $\boldsymbol{u}$ 求变分偏导数, 得到

$$\frac{\delta}{\delta\boldsymbol{u}}\left(\frac{\partial\mathcal{S}^1}{\partial t}\right)=\frac{\delta^2\mathcal{S}^1}{\delta\boldsymbol{u}\delta\widehat{\boldsymbol{p}}}\widehat{\boldsymbol{A}}^{-1}\left(\frac{\partial\widehat{\boldsymbol{A}}}{\partial t}\widehat{\boldsymbol{p}}\right)+\left(\frac{\delta^2\mathcal{S}^1}{\delta\boldsymbol{u}\delta\widehat{\boldsymbol{p}}}\widehat{\boldsymbol{A}}^{-1}+\boldsymbol{I}\right)\frac{\delta\mathcal{H}}{\delta\widehat{\boldsymbol{u}}}.\tag{7.2.36}$$

(7.2.35) 式加上 (7.2.36) 式, 变形整理, 得到

$$\boldsymbol{0}=\left(\frac{\delta^2\mathcal{S}^1}{\delta\boldsymbol{u}\delta\widehat{\boldsymbol{p}}}\widehat{\boldsymbol{A}}^{-1}+\boldsymbol{I}\right)\left(\widehat{\boldsymbol{A}}\frac{\partial\widehat{\boldsymbol{p}}}{\partial t}+\frac{\partial\widehat{\boldsymbol{A}}}{\partial t}\widehat{\boldsymbol{p}}+\frac{\delta\mathcal{H}}{\delta\widehat{\boldsymbol{u}}}\right).\tag{7.2.37}$$

对 (7.2.25) 式中的第一个关系式关于 $\boldsymbol{u}$ 求变分偏导数, 得到

$$\frac{\delta^2\mathcal{S}^1}{\delta\boldsymbol{u}\delta\widehat{\boldsymbol{p}}}=\widehat{\boldsymbol{A}}\frac{\delta\widehat{\boldsymbol{u}}}{\delta\boldsymbol{u}}-\widehat{\boldsymbol{A}}.\tag{7.2.38}$$

将上式代入 (7.2.37) 式, 再利用 $\dfrac{\delta\widehat{\boldsymbol{u}}}{\delta\boldsymbol{u}}$ 的可逆性, 则 (7.2.37) 式可改写为

$$\boldsymbol{0}=\widehat{\boldsymbol{A}}\frac{\partial\widehat{\boldsymbol{p}}}{\partial t}+\frac{\partial\widehat{\boldsymbol{A}}}{\partial t}\widehat{\boldsymbol{p}}+\frac{\delta\mathcal{H}}{\delta\widehat{\boldsymbol{u}}}.\tag{7.2.39}$$

将 (7.2.3) 式和 (7.2.12) 式代入上式, 则上式等价变形为 Birkhoff 系统 (7.2.2) 的第一个方程.

我们也可以通过和上面类似的过程由 (7.2.25) 式推出 Birkhoff 系统 (7.2.2) 的第二个方程: 首先, 对 (7.2.25) 式中的第一个关系式两端关于 t 求偏导数, 再对 (7.2.34) 式两端关于 $\widehat{\boldsymbol{p}}$ 求变分偏导数; 然后, 消掉 $\dfrac{\partial}{\partial t}\left(\dfrac{\delta \mathcal{S}^1}{\delta \widehat{\boldsymbol{p}}}\right)$, 再利用 Birkhoff 系统 (7.2.2) 的第一个方程就可推出它的第二个方程. 这说明生成泛函 $\mathcal{S}^1(\boldsymbol{u},\widehat{\boldsymbol{p}},t)$ 生成的映射 ϕ_t 是 Birkhoff 系统的解. □

下面我们讨论生成泛函 $\mathcal{S}^2(\widehat{\boldsymbol{u}},\boldsymbol{p},t)$ 和 $\mathcal{S}^3(\boldsymbol{v},\boldsymbol{w},t)$ 对应的 H-J 方程. 对 (7.2.29) 式两端关于 t 求偏导数, 得到

$$\int_x (\boldsymbol{A}\boldsymbol{p})^{\mathrm{T}}\frac{\partial \widehat{\boldsymbol{u}}}{\partial t}\mathrm{d}x-\frac{\partial \mathcal{S}^2}{\partial t}-\int_x\left(\frac{\partial \widehat{\boldsymbol{u}}}{\partial t}\right)^{\mathrm{T}}\frac{\delta \mathcal{S}^2}{\delta \widehat{\boldsymbol{u}}}\mathrm{d}x=\frac{\partial \mathcal{S}}{\partial t}+\int_x\left(\frac{\partial \widehat{\boldsymbol{u}}}{\partial t}\right)^{\mathrm{T}}\frac{\delta \mathcal{S}}{\delta \widehat{\boldsymbol{u}}}\mathrm{d}x. \tag{7.2.40}$$

又由 (7.2.27) 式得到 $\mathcal{S}^2(\widehat{\boldsymbol{u}},\boldsymbol{p},t)$ 的变分偏导数:

$$\frac{\delta \mathcal{S}^2}{\delta \widehat{\boldsymbol{u}}}=-\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}}+\boldsymbol{A}\boldsymbol{p},\quad \frac{\delta \mathcal{S}^2}{\delta \boldsymbol{p}}=\boldsymbol{A}(\widehat{\boldsymbol{u}}-\boldsymbol{u}). \tag{7.2.41}$$

将 (7.2.41) 式中的第一个关系式代入 (7.2.40) 式, 得到

$$\frac{\partial \mathcal{S}^2}{\partial t}=-\frac{\partial \mathcal{S}}{\partial t}. \tag{7.2.42}$$

同时 (7.2.41) 式意味着

$$\widehat{\boldsymbol{p}}=\widehat{\boldsymbol{A}}^{-1}\left(\boldsymbol{A}\boldsymbol{p}-\frac{\delta \mathcal{S}^2}{\delta \widehat{\boldsymbol{u}}}\right). \tag{7.2.43}$$

将 (7.2.42) 式和 (7.2.43) 式代入 (7.2.14) 式就可得到 $\mathcal{S}^2(\widehat{\boldsymbol{u}},\boldsymbol{p},t)$ 的 H-J 方程. 所以有下面的定理:

定理 7.2.4　假设生成泛函 $\mathcal{S}^2(\widehat{\boldsymbol{u}},\boldsymbol{p},t)$ 是 H-J 方程

$$\frac{\partial \mathcal{S}^2}{\partial t}=(\mathcal{B}+\overline{\mathcal{B}})\left(\widehat{\boldsymbol{u}},\widehat{\boldsymbol{A}}^{-1}\left(\boldsymbol{A}\boldsymbol{p}-\frac{\delta \mathcal{S}^2}{\delta \widehat{\boldsymbol{u}}}\right),t\right) \tag{7.2.44}$$

的解, 并且满足初始条件 $\mathcal{S}^2(\widehat{\boldsymbol{u}},\boldsymbol{p},t_0)=0$. 如果 $\dfrac{\delta \boldsymbol{u}}{\delta \widehat{\boldsymbol{u}}}$ 是可逆的, 则

(7.2.41) 式决定的映射 $\phi_t:(\boldsymbol{u},\boldsymbol{p})\mapsto(\widehat{\boldsymbol{u}},\widehat{\boldsymbol{p}})$ 是 Birkhoff 系统 (7.2.2) 的解.

以上定理的证明和定理 7.2.3 的证明过程是类似的, 这里从略.

注 7.2.2　尽管 $\mathcal{S}^2(\widehat{\boldsymbol{u}},\boldsymbol{p},t)$ 的 H-J 方程形式比 $\mathcal{S}^1(\boldsymbol{u},\widehat{\boldsymbol{p}},t)$ 的 H-J 方程形式紧凑, 但是 $\mathcal{S}^1(\widehat{\boldsymbol{u}},\boldsymbol{p},t)$ 和 $\mathcal{S}^2(\widehat{\boldsymbol{u}},\boldsymbol{p},t)$ 在构造保结构数值格式方面同样有效.

为了得到 $\mathcal{S}^3(\boldsymbol{v},\boldsymbol{w},t)$ 的 H-J 方程, 对 (7.2.31) 式两端关于 t 求偏导数并整理变形为

$$\begin{aligned}&\frac{\partial\mathcal{S}^3}{\partial t}+\int_x\left(\frac{\partial\boldsymbol{w}}{\partial t}\right)^{\mathrm{T}}\frac{\delta\mathcal{S}^3}{\delta\boldsymbol{w}}\mathrm{d}x+\int_x\left(\frac{\partial\boldsymbol{v}}{\partial t}\right)^{\mathrm{T}}\frac{\delta\mathcal{S}^3}{\delta\boldsymbol{v}}\mathrm{d}x\\&=\int_x\left(\frac{\partial\boldsymbol{w}}{\partial t}\right)^{\mathrm{T}}(\widehat{\boldsymbol{u}}-\boldsymbol{u})\mathrm{d}x+\int_x\boldsymbol{w}^{\mathrm{T}}\frac{\partial\widehat{\boldsymbol{u}}}{\partial t}\mathrm{d}x\\&\quad-\int_x\left(\frac{\partial\widehat{\boldsymbol{u}}}{\partial t}\right)^{\mathrm{T}}\frac{\delta\mathcal{S}}{\delta\widehat{\boldsymbol{u}}}\mathrm{d}x-\frac{\partial\mathcal{S}}{\partial t}.\end{aligned}\tag{7.2.45}$$

又由 (7.2.30) 式和 (7.2.31) 式得到 $\mathcal{S}^3(\boldsymbol{v},\boldsymbol{w},t)$ 的变分偏导数

$$\frac{\delta\mathcal{S}^3}{\delta\boldsymbol{w}}=\widehat{\boldsymbol{u}}-\boldsymbol{u},\quad-\frac{\delta\mathcal{S}^3}{\delta\boldsymbol{v}}=\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}}-\boldsymbol{A}\boldsymbol{p}.\tag{7.2.46}$$

方程组 (7.2.46) 等价于

$$\widehat{\boldsymbol{u}}=\boldsymbol{v}+\frac{1}{2}\frac{\delta\mathcal{S}^3}{\delta\boldsymbol{w}},\quad\widehat{\boldsymbol{p}}=\widehat{\boldsymbol{A}}^{-1}\left(\boldsymbol{w}-\frac{1}{2}\frac{\delta\mathcal{S}^3}{\delta\boldsymbol{v}}\right).\tag{7.2.47}$$

将变分偏导数 (7.2.46) 代入 (7.2.45) 式, 得到

$$\frac{\partial\mathcal{S}^3}{\partial t}=\int_x(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}})^{\mathrm{T}}\frac{\partial\widehat{\boldsymbol{u}}}{\partial t}\mathrm{d}x-\int_x\left(\frac{\partial\widehat{\boldsymbol{u}}}{\partial t}\right)^{\mathrm{T}}\frac{\delta\mathcal{S}}{\delta\widehat{\boldsymbol{u}}}\mathrm{d}x-\frac{\partial\mathcal{S}}{\partial t}.$$

因为 $\dfrac{\delta\mathcal{S}}{\delta\widehat{\boldsymbol{u}}}=\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}}$, 所以由上面的等式推出

$$\frac{\partial\mathcal{S}^3}{\partial t}=-\frac{\partial\mathcal{S}}{\partial t}.$$

将这个关系式和 (7.2.47) 式代入 H-J 方程 (7.2.14), 得到 $\mathcal{S}^3(\boldsymbol{v},\boldsymbol{w},t)$ 的 H-J 方程. 因此有下面的定理:

定理 7.2.5 假设 $\mathcal{S}^3(\boldsymbol{v},\boldsymbol{w},t)$ 是 H-J 方程

$$\frac{\partial\mathcal{S}^3}{\partial t}(\boldsymbol{v},\boldsymbol{w},t)=\mathcal{H}\left(\boldsymbol{v}+\frac{1}{2}\frac{\delta\mathcal{S}^3}{\delta\boldsymbol{w}},\widehat{\boldsymbol{A}}^{-1}\left(\boldsymbol{w}-\frac{1}{2}\frac{\delta\mathcal{S}^3}{\delta\boldsymbol{v}}\right),t\right)\tag{7.2.48}$$

的解, 并且满足初始条件 $\mathcal{S}^3(\boldsymbol{v},\boldsymbol{w},t_0)=0$. 如果矩阵

$$\begin{pmatrix}\boldsymbol{I}+\dfrac{1}{2}\dfrac{\delta^2\mathcal{S}^3}{\delta\boldsymbol{v}\delta\boldsymbol{w}} & \boldsymbol{0}\\ \boldsymbol{0} & -\dfrac{1}{2}\dfrac{\delta^2\mathcal{S}^3}{\delta\boldsymbol{v}^2}\end{pmatrix}\tag{7.2.49}$$

是可逆的, 则 (7.2.46) 式决定的映射 $\phi_t:(\boldsymbol{u},\boldsymbol{p})\mapsto(\widehat{\boldsymbol{u}},\widehat{\boldsymbol{p}})$ 是 Birkhoff 系统 (7.2.2) 的解.

证明 对 (7.2.46) 式中的第二个关系式两端关于 t 求偏导数, 得到

$$-\frac{\partial}{\partial t}\frac{\delta\mathcal{S}^3}{\delta\boldsymbol{v}}-\frac{1}{2}\frac{\delta^2\mathcal{S}^3}{\delta\boldsymbol{w}\delta\boldsymbol{v}}\frac{\partial(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}})}{\partial t}-\frac{1}{2}\frac{\delta^2\mathcal{S}^3}{\delta\boldsymbol{v}^2}\frac{\partial\widehat{\boldsymbol{u}}}{\partial t}=\frac{\partial(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}})}{\partial t}.\tag{7.2.50}$$

同时, 对 (7.2.48) 式两端关于 $\boldsymbol{v}$ 求变分偏导数, 得到

$$\frac{\delta}{\delta\boldsymbol{v}}\left(\frac{\partial\mathcal{S}^3}{\partial t}\right)=\left(\boldsymbol{I}+\frac{1}{2}\frac{\delta^2\mathcal{S}^3}{\delta\boldsymbol{v}\delta\boldsymbol{w}}\right)\frac{\delta\mathcal{H}}{\delta\widehat{\boldsymbol{u}}}-\frac{1}{2}\frac{\delta^2\mathcal{S}^3}{\delta\boldsymbol{v}^2}\widehat{\boldsymbol{A}}^{-1}\frac{\delta\mathcal{H}}{\delta\widehat{\boldsymbol{p}}}.\tag{7.2.51}$$

(7.2.50) 式与 (7.2.51) 式相加, 得到

$$\left(\boldsymbol{I}+\frac{1}{2}\frac{\delta^2\mathcal{S}^3}{\delta\boldsymbol{v}\delta\boldsymbol{w}}\right)\left(\frac{\partial(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}})}{\partial t}+\frac{\delta\mathcal{H}}{\delta\widehat{\boldsymbol{u}}}\right)=\frac{1}{2}\frac{\delta^2\mathcal{S}^3}{\delta\boldsymbol{v}^2}\left(\widehat{\boldsymbol{A}}^{-1}\frac{\delta\mathcal{H}}{\delta\widehat{\boldsymbol{p}}}-\frac{\partial\widehat{\boldsymbol{u}}}{\partial t}\right).\tag{7.2.52}$$

因为矩阵 (7.2.49) 是可逆的, 所以由上式推出

$$\frac{\partial(\widehat{\boldsymbol{A}}\widehat{\boldsymbol{p}})}{\partial t}+\frac{\delta\mathcal{H}}{\delta\widehat{\boldsymbol{u}}}=\boldsymbol{0},\quad \widehat{\boldsymbol{A}}^{-1}\frac{\delta\mathcal{H}}{\delta\widehat{\boldsymbol{p}}}-\frac{\partial\widehat{\boldsymbol{u}}}{\partial t}=\boldsymbol{0}.$$

再由 (7.2.3) 式和 (7.2.12) 式知道, 上面的方程组等价于 Birkhoff 系统 (7.2.2). □

7.2.4 一类特殊生成泛函 $\mathcal{S}^4$ 的 H-J 方程

如果上一小节的矩阵 $\widehat{\boldsymbol{A}}$ 可以这样分解: $\widehat{\boldsymbol{A}}=\widehat{\boldsymbol{M}}_1\cdot\widehat{\boldsymbol{M}}_2=\widehat{\boldsymbol{M}}_2\cdot\widehat{\boldsymbol{M}}_1$, $\widehat{\boldsymbol{M}}_1$ 和 $\widehat{\boldsymbol{M}}_2$ 在初始时刻 t_0 分别为 $\boldsymbol{M}_1$ 和 $\boldsymbol{M}_2$, 那么我们可以构造一个泛函 $\mathcal{S}^4(\boldsymbol{v},\boldsymbol{w},t)$, 使得

$$\frac{\delta\mathcal{S}^4}{\delta\boldsymbol{v}}=\boldsymbol{M}_2\boldsymbol{p}-\widehat{\boldsymbol{M}}_2\widehat{\boldsymbol{p}},\quad \frac{\delta\mathcal{S}^4}{\delta\boldsymbol{w}}=\widehat{\boldsymbol{M}}_1\widehat{\boldsymbol{u}}-\boldsymbol{M}_1\boldsymbol{u}, \tag{7.2.53}$$

其中

$$\boldsymbol{v}=(\widehat{\boldsymbol{M}}_1\widehat{\boldsymbol{u}}+\boldsymbol{M}_1\boldsymbol{u})/2,\quad \boldsymbol{w}=(\widehat{\boldsymbol{M}}_2\widehat{\boldsymbol{p}}+\boldsymbol{M}_2\boldsymbol{p})/2. \tag{7.2.54}$$

因此

$$\widehat{\boldsymbol{u}}=\widehat{\boldsymbol{M}}_1^{-1}\left(\boldsymbol{v}+\frac{1}{2}\frac{\delta\mathcal{S}^4}{\delta\boldsymbol{w}}\right),\quad \widehat{\boldsymbol{p}}=\widehat{\boldsymbol{M}}_2^{-1}\left(\boldsymbol{w}-\frac{1}{2}\frac{\delta\mathcal{S}^4}{\delta\boldsymbol{v}}\right), \tag{7.2.55}$$

并且有

$$\begin{pmatrix}\left(\dfrac{\delta\widehat{\boldsymbol{u}}}{\delta\boldsymbol{v}}\right)^{\mathrm{T}} & \left(\dfrac{\delta\widehat{\boldsymbol{p}}}{\delta\boldsymbol{v}}\right)^{\mathrm{T}}\\ \left(\dfrac{\delta\widehat{\boldsymbol{u}}}{\delta\boldsymbol{w}}\right)^{\mathrm{T}} & \left(\dfrac{\delta\widehat{\boldsymbol{p}}}{\delta\boldsymbol{w}}\right)^{\mathrm{T}}\end{pmatrix}
=\begin{pmatrix}\left(\boldsymbol{I}+\dfrac{1}{2}\dfrac{\delta^2\mathcal{S}^4}{\delta\boldsymbol{v}\delta\boldsymbol{w}}\right)^{\mathrm{T}}\widehat{\boldsymbol{M}}_1^{-\mathrm{T}} & \left(-\dfrac{1}{2}\dfrac{\delta^2\mathcal{S}^4}{\delta\boldsymbol{v}^2}\right)^{\mathrm{T}}\widehat{\boldsymbol{M}}_2^{-\mathrm{T}}\\ \left(\dfrac{1}{2}\dfrac{\delta^2\mathcal{S}^4}{\delta\boldsymbol{w}^2}\right)^{\mathrm{T}}\widehat{\boldsymbol{M}}_1^{-\mathrm{T}} & \left(\boldsymbol{I}-\dfrac{1}{2}\dfrac{\delta^2\mathcal{S}^4}{\delta\boldsymbol{w}\delta\boldsymbol{v}}\right)^{\mathrm{T}}\widehat{\boldsymbol{M}}_2^{-\mathrm{T}}\end{pmatrix}. \tag{7.2.56}$$

定理 7.2.6 假设 $\mathcal{S}^4(\boldsymbol{v},\boldsymbol{w},t)$ 是 H-J 方程

$$\frac{\partial\mathcal{S}^4}{\partial t}=(\mathcal{B}+\overline{\mathcal{B}})\left(\widehat{\boldsymbol{M}}_1^{-1}\left(\boldsymbol{v}+\frac{1}{2}\frac{\delta\mathcal{S}^4}{\delta\boldsymbol{w}}\right),\widehat{\boldsymbol{M}}_2^{-1}\left(\boldsymbol{w}-\frac{1}{2}\frac{\delta\mathcal{S}^4}{\delta\boldsymbol{v}}\right)\right) \tag{7.2.57}$$

的解, 并且满足初始条件 $\mathcal{S}^4(\boldsymbol{v},\boldsymbol{w},t_0)=0$, 其中 $\overline{\mathcal{B}}$ 是如下泛函:

$$\overline{\mathcal{B}}(\widehat{\boldsymbol{u}},\widehat{\boldsymbol{p}})=\frac{1}{2}\left(\widehat{\boldsymbol{u}}^{\mathrm{T}}\left(\frac{\partial\boldsymbol{F}_1}{\partial t}-\widehat{\boldsymbol{M}}_1\frac{\partial\widehat{\boldsymbol{M}}_2}{\partial t}\widehat{\boldsymbol{p}}\right)+\widehat{\boldsymbol{p}}^{\mathrm{T}}\left(\frac{\partial\boldsymbol{F}_2}{\partial t}+\widehat{\boldsymbol{M}}_2\frac{\partial\widehat{\boldsymbol{M}}_1}{\partial t}\widehat{\boldsymbol{u}}\right)\right), \tag{7.2.58}$$

且 $\overline{\mathcal{B}}$ 有变分偏导数

$$\frac{\delta\overline{\mathcal{B}}}{\delta\widehat{\boldsymbol{z}}}=\left(\frac{\partial\boldsymbol{F}_1}{\partial t}-\widehat{\boldsymbol{M}}_1\frac{\partial\widehat{\boldsymbol{M}}_2}{\partial t}\widehat{\boldsymbol{p}},\frac{\partial\boldsymbol{F}_2}{\partial t}+\widehat{\boldsymbol{M}}_2\frac{\partial\widehat{\boldsymbol{M}}_1}{\partial t}\widehat{\boldsymbol{u}}\right)^{\mathrm{T}}. \tag{7.2.59}$$

如果矩阵 (7.2.56) 是可逆的, 则 (7.2.53) 式决定的映射 $\phi_t:(\boldsymbol{u},\boldsymbol{p})\mapsto(\widehat{\boldsymbol{u}},\widehat{\boldsymbol{p}})$ 是 Birkhoff 系统 (7.2.2) 的解.

证明　对方程 (7.2.57) 两端关于 $\boldsymbol{v},\boldsymbol{w}$ 求变分偏导数, 得到

$$\begin{pmatrix}\dfrac{\delta}{\delta\boldsymbol{v}}\left(\dfrac{\partial\mathcal{S}^4}{\partial t}\right)\\ \dfrac{\delta}{\delta\boldsymbol{w}}\left(\dfrac{\partial\mathcal{S}^4}{\partial t}\right)\end{pmatrix}=\begin{pmatrix}\left(\dfrac{\delta\widehat{\boldsymbol{u}}}{\delta\boldsymbol{v}}\right)^{\mathrm{T}} & \left(\dfrac{\delta\widehat{\boldsymbol{p}}}{\delta\boldsymbol{v}}\right)^{\mathrm{T}}\\ \left(\dfrac{\delta\widehat{\boldsymbol{u}}}{\delta\boldsymbol{w}}\right)^{\mathrm{T}} & \left(\dfrac{\delta\widehat{\boldsymbol{p}}}{\delta\boldsymbol{w}}\right)^{\mathrm{T}}\end{pmatrix}\begin{pmatrix}\dfrac{\delta(\mathcal{B}+\overline{\mathcal{B}})}{\delta\widehat{\boldsymbol{u}}}\\ \dfrac{\delta(\mathcal{B}+\overline{\mathcal{B}})}{\delta\widehat{\boldsymbol{p}}}\end{pmatrix}. \tag{7.2.60}$$

再对 (7.2.53) 式中的两个关系式两端关于 t 求偏导数, 得到

$$\begin{aligned}&\begin{pmatrix}\dfrac{\partial}{\partial t}\left(\dfrac{\delta\mathcal{S}^4}{\delta\boldsymbol{v}}\right)\\ \dfrac{\partial}{\partial t}\left(\dfrac{\delta\mathcal{S}^4}{\delta\boldsymbol{w}}\right)\end{pmatrix}\\ &=\begin{pmatrix}\left(\boldsymbol{I}+\dfrac{1}{2}\dfrac{\delta^2\mathcal{S}^4}{\delta\boldsymbol{w}\delta\boldsymbol{v}}\right)^{\mathrm{T}}\widehat{\boldsymbol{M}}_1^{-\mathrm{T}} & \left(-\dfrac{1}{2}\dfrac{\delta^2\mathcal{S}^4}{\delta\boldsymbol{v}^2}\right)^{\mathrm{T}}\widehat{\boldsymbol{M}}_2^{-\mathrm{T}}\\ \left(\dfrac{1}{2}\dfrac{\delta^2\mathcal{S}^4}{\delta\boldsymbol{w}^2}\right)^{\mathrm{T}}\widehat{\boldsymbol{M}}_1^{-\mathrm{T}} & \left(\boldsymbol{I}-\dfrac{1}{2}\dfrac{\delta^2\mathcal{S}^4}{\delta\boldsymbol{v}\delta\boldsymbol{w}}\right)^{\mathrm{T}}\widehat{\boldsymbol{M}}_2^{-\mathrm{T}}\end{pmatrix}\\ &\quad\cdot\begin{pmatrix}-\widehat{\boldsymbol{M}}_1\widehat{\boldsymbol{M}}_2\dfrac{\partial\widehat{\boldsymbol{p}}}{\partial t}-\widehat{\boldsymbol{M}}_1\dfrac{\partial\widehat{\boldsymbol{M}}_2}{\partial t}\widehat{\boldsymbol{p}}\\ \widehat{\boldsymbol{M}}_2\widehat{\boldsymbol{M}}_1\dfrac{\partial\widehat{\boldsymbol{u}}}{\partial t}+\widehat{\boldsymbol{M}}_2\dfrac{\partial\widehat{\boldsymbol{M}}_1}{\partial t}\widehat{\boldsymbol{u}}\end{pmatrix}.\end{aligned} \tag{7.2.61}$$

对比 (7.2.61) 式和 (7.2.60) 式, 再利用 (7.2.56) 式, 我们得到

$$\begin{pmatrix}-\widehat{\boldsymbol{M}}_1\widehat{\boldsymbol{M}}_2\dfrac{\partial\widehat{\boldsymbol{p}}}{\partial t}-\widehat{\boldsymbol{M}}_1\dfrac{\partial\widehat{\boldsymbol{M}}_2}{\partial t}\widehat{\boldsymbol{p}}\\ \widehat{\boldsymbol{M}}_2\widehat{\boldsymbol{M}}_1\dfrac{\partial\widehat{\boldsymbol{u}}}{\partial t}+\widehat{\boldsymbol{M}}_2\dfrac{\partial\widehat{\boldsymbol{M}}_1}{\partial t}\widehat{\boldsymbol{u}}\end{pmatrix}=\begin{pmatrix}\dfrac{\delta(\mathcal{B}+\overline{\mathcal{B}})}{\delta\widehat{\boldsymbol{u}}}\\ \dfrac{\delta(\mathcal{B}+\overline{\mathcal{B}})}{\delta\widehat{\boldsymbol{p}}}\end{pmatrix}. \tag{7.2.62}$$

这正是 Birkhoff 系统 (7.2.2). □

注 7.2.3　若 $\boldsymbol{M}_1=\boldsymbol{M}_2=\boldsymbol{M}$, 则可以选择 $\boldsymbol{w}=(\widehat{\boldsymbol{M}}\widehat{\boldsymbol{p}}+\boldsymbol{M}p)/2$ 和 $\boldsymbol{v}=(\widehat{\boldsymbol{M}}\widehat{\boldsymbol{u}}+\boldsymbol{M}\boldsymbol{u})/2$ 作为泛函 $\mathcal{S}^4(\boldsymbol{v},\boldsymbol{w},t)$ 的变量. 此时, $\overline{\mathcal{B}}=0$, 并且基于这样的 $\mathcal{S}^4(\boldsymbol{v},\boldsymbol{w},t)$ 构造的数值格式是对称的.

§7.3 基于 $\mathcal{S}^2, \mathcal{S}^3$ 和 $\mathcal{S}^4$ 的半离散数值格式

在本节中, 我们将以 Birkhoff 系统 (7.1.1) 为例说明如何基于生成泛函构造保 Birkhoff 辛结构的数值算法. 因为 Birkhoff 辛结构和经典 Hamilton 辛结构的不同, Hamilton 辛结构是 $\boldsymbol{J}$-辛结构, 而 Birkhoff 辛结构是 $\boldsymbol{K}(\boldsymbol{z},x,t)$-辛结构, 前者是唯一的, 后者因为 $\boldsymbol{K}(\boldsymbol{z},x,t)$ 不同而不同, 所以很难给出一个系统的、适用于所有 Birkhoff 系统的构造数值格式的方法. 但是, 因为我们在上一节已经详细讨论了生成泛函和一般 Birkhoff 辛结构的关系, 所以只要掌握了本节的方法, 就可以很容易地把这个方法推广应用到其他 Birkhoff 系统.

对于 Birkhoff 系统 (7.1.1), 有 $\widehat{A}=\mathrm{e}^{\lambda t}$. 不妨令初始时刻为 $t_0=0$, 则对应有

$$A=\mathrm{e}^0=1,\quad \widehat{A}=\mathrm{e}^{\lambda t}, \tag{7.3.1}$$

对应的泛函 $\mathcal{H}(\widehat{u},\widehat{p},t)=\mathcal{B}+\overline{\mathcal{B}}$ 为

$$\mathcal{H}=\frac{1}{2}\mathrm{e}^{\lambda t}\int_V\left[\widehat{u}^2+c(\widehat{u}_x^2+\widehat{u}_y^2)+\widehat{p}^2\right]\mathrm{d}V, \tag{7.3.2}$$

其中 V 代表 (x,y) 空间区域, $\mathcal{B}$ 和 $\overline{\mathcal{B}}$ 分别为

$$\begin{aligned}\mathcal{B}&=\frac{1}{2}\mathrm{e}^{\lambda t}\int_V\left[\widehat{u}^2+c(\widehat{u}_x^2+\widehat{u}_y^2)+\widehat{p}^2+\lambda\widehat{u}\widehat{p}\right]\mathrm{d}V,\\ \overline{\mathcal{B}}&=-\frac{1}{2}\mathrm{e}^{\lambda t}\int_V\lambda\widehat{p}\widehat{u}\mathrm{d}V.\end{aligned} \tag{7.3.3}$$

7.3.1 基于 $\mathcal{S}^2$ 的半离散数值格式

H-J 方程 (7.2.44) 在本例中具体表示为

$$\begin{aligned}\frac{\partial\mathcal{S}^2}{\partial t}=\frac{1}{2}\int_V\Bigg[&\mathrm{e}^{\lambda t}\left(\widehat{u}^2+c(\widehat{u}_x^2+\widehat{u}_y^2)\right)\\ &+\mathrm{e}^{-\lambda t}\left(p^2-2p\frac{\delta\mathcal{S}^2}{\delta\widehat{u}}+\left(\frac{\delta\mathcal{S}^2}{\delta\widehat{u}}\right)^2\right)\Bigg]\mathrm{d}V.\end{aligned} \tag{7.3.4}$$

我们假设上面的方程有如下级数形式的解:

$$\mathcal{S}^2(\widehat{u},p,t)=t\mathcal{G}_1(\widehat{u},p)+t^2\mathcal{G}_2(\widehat{u},p)+t^3\mathcal{G}_3(\widehat{u},p)+\cdots,\tag{7.3.5}$$

其中 $\mathcal{G}_i(i=1,2,\cdots)$ 都是泛函. 把这个形式解代入方程 (7.3.4), 两端都整理成 t 的级数形式. 下面列出方程右端的不同次数项系数, 我们用 $\mathcal{G}_i'$ 表示 $\dfrac{\delta\mathcal{G}_i}{\delta\widehat{u}}$:

$$\begin{aligned}
&t^0:\frac{1}{2}\int_V\left[\widehat{u}^2+c\big(\widehat{u}_x^2+\widehat{u}_y^2\big)+p^2\right]\mathrm{d}V,\\
&t^1:\frac{1}{2}\int_V\left[\lambda\big(\widehat{u}^2+c\big(\widehat{u}_x^2+\widehat{u}_y^2\big)-p^2\big)-2p\mathcal{G}_1'\right]\mathrm{d}V,\\
&t^2:\frac{1}{2}\int_V\left[\frac{\lambda^2}{2}\big(\widehat{u}^2+c\big(\widehat{u}_x^2+\widehat{u}_y^2\big)+p^2\big)+\mathcal{G}_1'^2-2p\mathcal{G}_2'+2\lambda p\mathcal{G}_1'\right]\mathrm{d}V,\\
&t^3:\frac{1}{2}\int_V\left[\frac{\lambda^3}{6}\big(\widehat{u}^2+c\big(\widehat{u}_x^2+\widehat{u}_y^2\big)-p^2\big)+2\mathcal{G}_1'\mathcal{G}_2'-2p\mathcal{G}_3'\right.\\
&\qquad\left.-\lambda\mathcal{G}_1'^2+2\lambda p\mathcal{G}_2'-\lambda^2p\mathcal{G}_1'\right]\mathrm{d}V,\\
&\cdots\cdots
\end{aligned}$$

方程 (7.3.4) 两端同类项的系数相等, 由此得到

$$\begin{aligned}
\mathcal{G}_1=&\int_V\frac{1}{2}\left[\widehat{u}^2+c\big(\widehat{u}_x^2+\widehat{u}_y^2\big)+p^2\right]\mathrm{d}V,\\
\mathcal{G}_2=&\int_V\frac{1}{4}\left[\lambda\big(\widehat{u}^2+c\big(\widehat{u}_x^2+\widehat{u}_y^2\big)-p^2\big)-2p\big(\widehat{u}-c\big(\widehat{u}_{xx}+\widehat{u}_{yy}\big)\big)\right]\mathrm{d}V,\\
\mathcal{G}_3=&\int_V\frac{1}{6}\left[\frac{\lambda^2}{2}\big(\widehat{u}^2+c\big(\widehat{u}_x^2+\widehat{u}_y^2\big)+p^2\big)+\big(\widehat{u}-c\big(\widehat{u}_{xx}+\widehat{u}_{yy}\big)\big)^2\right.\\
&\left.+\lambda p\big(\widehat{u}-c\big(\widehat{u}_{xx}+\widehat{u}_{yy}\big)\big)+p^2\right]\mathrm{d}V,\\
&\cdots\cdots
\end{aligned}$$

原则上我们可以得到级数 $\mathcal{S}^2(\widehat{u},p,t)$ 的任意阶项系数 $\mathcal{G}_i$. 我们取 $\mathcal{S}^2(\widehat{u},p,t)$ 的 2 阶截断近似:

$$\mathcal{S}^2(\widehat{u},p,t)\approx t\mathcal{G}_1(\widehat{u},p)+t^2\mathcal{G}_2(\widehat{u},p).\tag{7.3.6}$$

这是 H-J 方程 (7.3.4) 的一个近似解, 所以它决定的映射是 Birkhoff 系统 (7.1.1) 的一个近似解. 具体做法为: 将 (7.3.6) 式代入方程组 (7.2.41), 得到

$$\begin{aligned} &\frac{2t+\lambda t^2}{2}\left(\widehat{u}-c(\widehat{u}_{xx}+\widehat{u}_{yy})\right)-\frac{t^2}{2}\left(p-c(p_{xx}+p_{yy})\right)=p-\mathrm{e}^{\lambda t}\widehat{p}, \\ &\frac{2t-\lambda t^2}{2}p-\frac{t^2}{2}\left(\widehat{u}-c(\widehat{u}_{xx}+\widehat{u}_{yy})\right)=\widehat{u}-u. \end{aligned} \tag{7.3.7}$$

这就是 Birkhoff 系统 (7.1.1) 的一个时间离散而空间连续的半离散形式.

定理 7.3.1 半离散方程组 (7.3.7) 所确定的半离散格式保持 Birkhoff 辛结构 $\int_V \mathrm{e}^{\lambda t}\delta\widehat{p}\wedge\delta\widehat{u}$, 且在时间方向是 2 阶精度.

证明 将方程组 (7.3.7) 中两方程的两端乘以 $\delta\widehat{u}$, 再关于空间区域积分, 得

$$\begin{aligned} &\int_V\left[\frac{2t+\lambda t^2}{2}\left(\widehat{u}-c(\widehat{u}_{xx}+\widehat{u}_{yy})\right)-\frac{t^2}{2}\left(p-c(p_{xx}+p_{yy})\right)\right]\delta\widehat{u} \\ &\quad=\int_V\left(p-\mathrm{e}^{\lambda t}\widehat{p}\right)\delta\widehat{u}, \\ &\int_V\left[\frac{2t-\lambda t^2}{2}p-\frac{t^2}{2}\left(\widehat{u}-c(\widehat{u}_{xx}+\widehat{u}_{yy})\right)\right]\delta p=\int_V(\widehat{u}-u)\delta p. \end{aligned} \tag{7.3.8}$$

对方程组 (7.3.8) 中两方程的两端变分, 整理得到

$$\begin{aligned} \int_V \mathrm{e}^{\lambda t}\delta\widehat{p}\wedge\delta\widehat{u}=&\int_V\frac{2t+\lambda t^2}{2}c\delta(\widehat{u}_{xx}+\widehat{u}_{yy})\wedge\delta\widehat{u} \\ &-\int_V\frac{ct^2}{2}\delta(p_{xx}+p_{yy})\wedge\delta\widehat{u}+\int_V\frac{2+t^2}{2}\delta p\wedge\delta\widehat{u} \end{aligned} \tag{7.3.9}$$

和

$$\int_V\delta p\wedge\delta u=\int_V\frac{2+t^2}{2}\delta p\wedge\delta\widehat{u}+\int_V\frac{ct^2}{2}\delta(\widehat{u}_{xx}+\widehat{u}_{yy})\wedge\delta p. \tag{7.3.10}$$

利用注 7.1.1 所假设的边界条件, 我们得到

$$\int_V\delta(\widehat{u}_{xx}+\widehat{u}_{yy})\wedge\delta\widehat{u}=-\int_V(\delta\widehat{u}_x\wedge\delta\widehat{u}_x+\delta\widehat{u}_y\wedge\delta\widehat{u}_y)=0 \tag{7.3.11}$$

和

$$\int_V \delta(\widehat{u}_{xx}+\widehat{u}_{yy})\wedge\delta p=\int_V \delta\widehat{u}\wedge\delta(p_{xx}+p_{yy}). \tag{7.3.12}$$

(7.3.10) 式减去 (7.3.9) 式, 同时利用 (7.3.11) 式和 (7.3.12) 式, 得到

$$\int_V \mathrm{e}^{\lambda t}\delta\widehat{p}\wedge\delta\widehat{u}=\int_V \delta p\wedge\delta u, \tag{7.3.13}$$

即格式 (7.3.8) 保持辛结构 $\int_V \mathrm{e}^{\lambda t}\delta\widehat{p}\wedge\delta\widehat{u}$.

下面我们证明格式是 2 阶精度的. 为了简便起见, 令 $z^{(2)}$ 和 $z^{(4)}$ 分别表示 $z_{xx}+z_{yy}$ 和 $z_{xxxx}+z_{yyyy}$, 其中 z 是 u 或者 p. 方程 (7.1.1) 的精确解可以展开为下面的级数形式:

$$\begin{aligned}\widehat{u}=&u+tp+\frac{t^2}{2}\big[-\big(u-cu^{(2)}\big)-\lambda p\big]\\&+\frac{t^3}{6}\big[\lambda\big(u-cu^{(2)}\big)-\big(p-cp^{(2)}\big)+\lambda^2p\big]+o(t^3),\\\widehat{p}=&p+t\big[-\big(u-cu^{(2)}\big)-\lambda p\big]+\frac{t^2}{2}\big[\lambda\big(u-cu^{(2)}\big)-\big(p-cp^{(2)}\big)+\lambda^2p\big]\\&+\frac{t^3}{6}\big[\big(1-\lambda^2\big)\big(u-cu^{(2)}\big)+2\lambda\big(p-cp^{(2)}\big)\\&-c\big(u^{(2)}-cu^{(4)}\big)-\lambda^3p\big]+o(t^3).\end{aligned}$$

而半离散方程组 (7.3.7) 有下面的级数形式解:

$$\begin{aligned}\widehat{u}=&u+tp+\frac{t^2}{2}\big[-\big(u-cu^{(2)}\big)-\lambda p\big]-\frac{t^3}{2}\big(p-cp^{(2)}\big)+o(t^3),\\\widehat{p}=&p+t\big[-\big(u-cu^{(2)}\big)-\lambda p\big]+\frac{t^2}{2}\big[\lambda\big(u-cu^{(2)}\big)-\big(p-cp^{(2)}\big)+\lambda^2p\big]\\&+\frac{t^3}{6}\big[3\big(u-cu^{(2)}\big)+3\lambda\big(p-cp^{(2)}\big)-3c\big(u^{(2)}-cu^{(4)}\big)-\lambda^3p\big]+o(t^3).\end{aligned}$$

对比上面两组级数形式解的系数, 可以看出格式 (7.3.7) 在时间方向具有 2 阶精度. □

7.3.2 基于 $\mathcal{S}^3$ 的半离散数值格式

在此例中, 泛函 $\mathcal{S}^3(v,w,t)$ 的 H-J 方程 (7.2.47) 具体为

$$\frac{\partial \mathcal{S}^3}{\partial t}=\frac{1}{2}\int_V\left\{\mathrm{e}^{\lambda t}\left[\widehat{u}^2+c(\widehat{u}_x^2+\widehat{u}_y^2)\right]+\mathrm{e}^{-\lambda t}\left[w^2-w\frac{\delta\mathcal{S}^3}{\delta v}+\frac{1}{4}\left(\frac{\delta\mathcal{S}^3}{\delta v}\right)^2\right]\right\}\mathrm{d}V, \tag{7.3.14}$$

其中

$$\begin{aligned}
\widehat{u}^2&=v^2+v\frac{\delta\mathcal{S}^3}{\delta w}+\frac{1}{4}\left(\frac{\delta\mathcal{S}^3}{\delta w}\right)^2,\\
\widehat{u}_x^2&=v_x^2+v_x\left(\frac{\delta\mathcal{S}^3}{\delta w}\right)_x+\frac{1}{4}\left(\frac{\delta\mathcal{S}^3}{\delta w}\right)_x^2,\\
\widehat{u}_y^2&=v_y^2+v_y\left(\frac{\delta\mathcal{S}^3}{\delta w}\right)_y+\frac{1}{4}\left(\frac{\delta\mathcal{S}^3}{\delta w}\right)_y^2.
\end{aligned}$$

同样, 假设 $\mathcal{S}^3(v,w,t)$ 有级数形式

$$\mathcal{S}^3(v,w,t)=t\mathcal{G}_1(v,w)+t^2\mathcal{G}_2(v,w)+t^3\mathcal{G}_3(v,w)+\cdots,$$

用 $\mathcal{G}_{iv}$ 和 $\mathcal{G}_{iw}$ 分别表示 $\dfrac{\delta\mathcal{G}_i}{\delta v}$ 和 $\dfrac{\delta\mathcal{G}_i}{\delta w}$. 将上述级数解代入方程 (7.3.14), 得到方程右端各次数项系数为

$$\begin{aligned}
t^0:&\ \frac{1}{2}\int_V\left[v^2+c\big(v_x^2+v_y^2\big)+w^2\right]\mathrm{d}V,\\
t^1:&\ \frac{1}{2}\int_V\Big[v\mathcal{G}_{1w}+cv_x\mathcal{G}_{1wx}+cv_y\mathcal{G}_{1wy}\\
&\quad+\lambda\big(v^2+c\big(v_x^2+v_y^2\big)-w^2\big)-w\mathcal{G}_{1v}\Big]\mathrm{d}V,\\
t^2:&\ \frac{1}{2}\int_V\Big[\frac{\lambda^2}{2}\big(v^2+c\big(v_x^2+v_y^2\big)+w^2\big)\\
&\quad+\lambda\big(v\mathcal{G}_{1w}+c\big(v_x\mathcal{G}_{1wx}+v_y\mathcal{G}_{1wy}\big)+w\mathcal{G}_{1v}\big)\\
&\quad+\big(v\mathcal{G}_{2w}+c\big(v_x\mathcal{G}_{2wx}+v_y\mathcal{G}_{2wy}\big)-w\mathcal{G}_{2v}\big)\\
&\quad+\frac{1}{4}\big(\mathcal{G}_{1w}^2+c\big(\mathcal{G}_{1wx}^2+\mathcal{G}_{1wy}^2\big)+\mathcal{G}_{1v}^2\big)\Big]\mathrm{d}V,
\end{aligned}$$

……

和 (7.3.14) 式左端同类项比较系数, 得到

$$
\begin{aligned}
\mathcal{G}_1 &= \frac{1}{2}\int_V \left[v^2 + c\left(v_x^2 + v_y^2\right) + w^2\right]\mathrm{d}V, \\
\mathcal{G}_2 &= \frac{1}{4}\int_V \left[\lambda\left(v^2 + c\left(v_x^2 + v_y^2\right) - w^2\right)\right]\mathrm{d}V, \\
&\cdots\cdots
\end{aligned}
$$

将截断形式 $\mathcal{S}^3 = t\mathcal{G}_1 + t^2\mathcal{G}_2$ 代入方程组 (7.2.46), 整理得到半离散格式

$$
\frac{2t - \lambda t^2}{2} w = \widehat{u} - u, \quad \frac{2t + \lambda t^2}{2}\left(v - c\left(v_{xx} + v_{yy}\right)\right) = p - \mathrm{e}^{\lambda t}\widehat{p}. \quad (7.3.15)
$$

该格式有如下级数展开式:

$$
\begin{aligned}
\widehat{u} &= u + tp + \frac{t^2}{2}\left[-\left(u - cu^{(2)}\right) - \lambda p\right] \\
&\quad + \frac{t^3}{6}\left[-\frac{3\lambda}{2}\left(u - cu^{(2)}\right) - \frac{3}{2}\left(p - cp^{(2)}\right) - \frac{3\lambda^2}{2}p\right] + o\left(t^3\right), \\
\widehat{p} &= p + t\left[-\left(u - cu^{(2)}\right) - \lambda p\right] + \frac{t^2}{2}\left[\lambda\left(u - cu^{(2)}\right) - \left(p - cp^{(2)}\right) + \lambda^2 p\right] \\
&\quad + \frac{t^3}{6}\left[\frac{3}{2}\left(u - cu^{(2)}\right) + 3\lambda\left(p - cp^{(2)}\right) - \frac{3c}{2}\left(u^{(2)} - cu^{(4)}\right) - \lambda^3 p\right] + o\left(t^3\right).
\end{aligned}
$$

和精确解的级数形式作对比, 显示该格式在时间方向上为 2 阶精度.

定理 7.3.2　*半离散格式 (7.3.15) 保持 Birkhoff 辛结构* $\int_V \mathrm{e}^{\lambda t}\delta\widehat{p}\wedge\delta\widehat{u}$, *且在时间方向上具有 2 阶精度.*

定理的证明过程类似于定理 7.3.1 的证明, 从略.

7.3.3　基于 $\mathcal{S}^4$ 的半离散数值格式

$\widehat{A}(t) = \mathrm{e}^{\lambda t}$ 可以分解为 $\mathrm{e}^{\frac{\lambda t}{2}} \cdot \mathrm{e}^{\frac{\lambda t}{2}}$. 令 v, w 分别为 $\mathcal{S}^4$ 的变量:

$$
v = \left(\mathrm{e}^{\frac{\lambda t}{2}}\widehat{u} + \mathrm{e}^{\frac{\lambda t_0}{2}}u\right)/2, \quad w = \left(\mathrm{e}^{\frac{\lambda t}{2}}\widehat{p} + \mathrm{e}^{\frac{\lambda t_0}{2}}p\right)/2,
$$

则对应的 H-J 方程为

$$
\frac{\partial \mathcal{S}^4}{\partial t} = \mathcal{B}\left(\widehat{u}, \widehat{p}, t\right) = \frac{1}{2}\mathrm{e}^{\lambda t}\int_V \left[\widehat{u}^2 + c\left(\widehat{u}_x^2 + \widehat{u}_y^2\right) + \widehat{p}^2 + \lambda\widehat{u}\widehat{p}\right]\mathrm{d}V, \quad (7.3.16)
$$

其中

$$\widehat{u} = \mathrm{e}^{-\frac{\lambda t}{2}}\left(v + \frac{1}{2}\frac{\delta\mathcal{S}^4}{\delta w}\right), \quad \widehat{p} = \mathrm{e}^{-\frac{\lambda t}{2}}\left(w - \frac{1}{2}\frac{\delta\mathcal{S}^4}{\delta v}\right).$$

假设 $S^4(v,w,t)$ 的级数展开式为

$$\mathcal{S}^4(v,w,t) = t\mathcal{G}_1 + t^2\mathcal{G}_2 + t^3\mathcal{G}_3 + \cdots,$$

$\mathcal{B}$ 的级数形式为

$$\mathcal{B} = \mathcal{B}_0(v,w) + t\mathcal{B}_1(v,w) + t^2\mathcal{B}_2(v,w) + \cdots, \tag{7.3.17}$$

其中

$$\begin{aligned}
\mathcal{B}_0 &= \frac{1}{2}\int_V \left[v^2 + c\left(v_x^2 + v_y^2\right) + w^2 + \lambda v w\right]\mathrm{d}V, \\
\mathcal{B}_1 &= \frac{1}{2}\int_V \left[v\mathcal{G}_{1w} + c\left(v_x\mathcal{G}_{1wx} + v_y\mathcal{G}_{1wy}\right) - w\mathcal{G}_{1v} - \lambda\frac{1}{2}\left(v\mathcal{G}_{1v} - w\mathcal{G}_{1w}\right)\right]\mathrm{d}V, \\
\mathcal{B}_2 &= \frac{1}{2}\int_V \left[v\mathcal{G}_{2w} + c\left(v_x\mathcal{G}_{2wx} + v_y\mathcal{G}_{2wy}\right) - w\mathcal{G}_{2v} - \lambda\frac{1}{2}\left(v\mathcal{G}_{2v} - w\mathcal{G}_{2w}\right)\right]\mathrm{d}V \\
&\quad + \frac{1}{8}\int_V \left[\mathcal{G}_{1w}^2 + c\left(\mathcal{G}_{1wx}^2 + \mathcal{G}_{1wy}^2\right) + \mathcal{G}_{1v}^2 - \lambda\mathcal{G}_{1v}\mathcal{G}_{1w}\right]\mathrm{d}V,
\end{aligned}$$

……

取 $\mathcal{S}^4$ 的 2 阶截断近似:

$$S^4 \approx t\mathcal{G}_1 + t^2\mathcal{G}_2,$$

其对应系数为

$$\mathcal{G}_1 = \frac{1}{2}\int_V \left[v^2 + c\left(v_x^2 + v_y^2\right) + w^2 + \lambda v w\right]\mathrm{d}V, \quad \mathcal{G}_2 = 0, \tag{7.3.18}$$

将 $\mathcal{S}^4$ 的截断形式代入方程组 (7.2.53), 得到半离散格式

$$\begin{aligned}
&\frac{t}{2}\left[\mathrm{e}^{\frac{\lambda t}{2}}\widehat{u} + u - c\left(\mathrm{e}^{\frac{\lambda t}{2}}\widehat{u}_{xx} + u_{xx} + \mathrm{e}^{\frac{\lambda t}{2}}\widehat{u}_{yy} + u_{yy}\right)\right] \\
&\quad + \frac{\lambda t}{4}\left(\mathrm{e}^{\frac{\lambda t}{2}}\widehat{p} + p\right) = p - \mathrm{e}^{\frac{\lambda t}{2}}\widehat{p}, \\
&\frac{t}{2}\left(\mathrm{e}^{\frac{\lambda t}{2}}\widehat{p} + p\right) + \frac{\lambda t}{4}\left(\mathrm{e}^{\frac{\lambda t}{2}}\widehat{u} + u\right) = \mathrm{e}^{\frac{\lambda t}{2}}\widehat{u} - u.
\end{aligned} \tag{7.3.19}$$

该格式有如下级数展开式:

$$
\begin{aligned}
\widehat{u} &= u + tp + \frac{t^2}{2}\big[-\big(u - cu^{(2)}\big) - \lambda p\big] \\
&\quad + \frac{t^3}{6}\left[\frac{\lambda^3}{16}u - \frac{9\lambda}{4}\big(u - cu^{(2)}\big) + \frac{3}{2}\big(p - cp^{(2)}\big) - \frac{15\lambda^2}{8}p\right] + o(t^3), \\
\widehat{p} &= p + t\big[-\big(u - cu^{(2)}\big) - \lambda p\big] + \frac{t^2}{2}\big[\lambda\big(u - cu^{(2)}\big) - \big(p - cp^{(2)}\big) + \lambda^2 p\big] \\
&\quad + \frac{t^3}{6}\left[\frac{12 - 3\lambda^2}{8}\big(u - cu^{(2)}\big) + \frac{9\lambda}{4}\big(p - cp^{(2)}\big)\right. \\
&\quad \left. - \frac{3c}{2}\big(u^{(2)} - cu^{(4)}\big) - \frac{7\lambda^3}{16}p\right] + o(t^3).
\end{aligned}
$$

定理 7.3.3 半离散格式 (7.3.19) 保持 Birkhoff 辛结构 $\int_V e^{\lambda t}\delta\widehat{p}\wedge\delta\widehat{u}$, 且在时间方向为 2 阶精度.

同样, 定理的证明类似于定理 7.3.1 的证明, 从略.

§7.4 数 值 实 验

本节将以二维声波方程为例, 对比无限维 Birkhoff 算法和一般算法的优劣.

7.4.1 声波方程

我们对声波方程 (7.1.1) 的无限维 Birkhoff 形式 (7.1.3) 进行离散: 首先时间方向分别采用 Birkhoff 算法 (7.3.7), (7.3.15) 和 (7.3.19) 离散, 然后在空间方向一律采用中点格式离散. 这样得到的三个全离散格式分别记为 S2, S3 和 S4. 我们称这三个格式为 **Birkhoff 辛格式**.

另外, 我们在时间和空间均采用中点格式离散, 得到一个全离散格式. 我们知道这种格式是 Hamilton 辛格式, 同时还是多辛格式, 我们用 “Hamiltonian” 记这个格式.

用 n 表示时间节点, (j,k) 表示空间节点, τ 表示时间步长, h_x 和 h_y 分别表示 x 和 y 方向的步长.

上述三个 Birkhoff 辛格式和 Hamiltonian 格式表示如下:

S2 格式:

$$
\begin{aligned}
& a_1 u_{j,k}^{n+1} - \frac{m_1}{h_x^2}\left(u_{j+1,k}^{n+1} + u_{j-1,k}^{n+1}\right) - \frac{m_1}{h_y^2}\left(u_{j,k+1}^{n+1} + u_{j,k-1}^{n+1}\right) \\
& = \frac{2\tau - \lambda\tau^2}{2} p_{j,k}^n + u_{j,k}^n, \\
& \mathrm{e}^{\lambda\tau} p_{j,k}^{n+1} = a_1 p_{j,k}^n - \left(b_1 + \frac{2b_1 c}{h_x^2} + \frac{2b_1 c}{h_y^2}\right) u_{j,k}^{n+1} \\
& \qquad + \frac{b_1 c}{h_x^2}\left(u_{j+1,k}^{n+1} + u_{j-1,k}^{n+1}\right) - \frac{b_1 c}{h_y^2}\left(u_{j,k+1}^{n+1} + u_{j,k-1}^{n+1}\right),
\end{aligned} \tag{7.4.1}
$$

其中

$$
m_1 = \frac{\tau^2}{2}c, \quad a_1 = \frac{2+\tau^2}{2} + \frac{2m_1}{h_x^2} + \frac{2m_1}{h_y^2}, \quad b_1 = \frac{2\tau + \lambda\tau^2}{2}.
$$

S3 格式:

$$
\begin{aligned}
& a_2 u_{j,k}^{n+1} - \frac{m_2}{h_x^2}\left(u_{j+1,k}^{n+1} + u_{j-1,k}^{n+1}\right) - \frac{m_2}{h_y^2}\left(u_{j,k+1}^{n+1} + u_{j,k-1}^{n+1}\right) \\
& = 2p_{j,k}^n - b_2 u_{j,k}^n + \frac{m_2}{h_x^2}\left(u_{j+1,k}^n + u_{j-1,k}^n\right) + \frac{m_2}{h_y^2}\left(u_{j,k+1}^n + u_{j,k-1}^n\right), \\
& \mathrm{e}^{\lambda\tau} p_{j,k}^{n+1} + \left[d_2 u_{j,k}^{n+1} - \frac{m_2}{h_x^2}\left(u_{j+1,k}^{n+1} + u_{j-1,k}^{n+1}\right) - \frac{m_2}{h_y^2}\left(u_{j,k+1}^{n+1} + u_{j,k-1}^{n+1}\right)\right] \\
& = p_{j,k}^n - \left[d_2 u_{j,k}^n - \frac{m_2}{h_x^2}\left(u_{j+1,k}^n + u_{j-1,k}^n\right) - \frac{m_2}{h_y^2}\left(u_{j,k+1}^n + u_{j,k-1}^n\right)\right],
\end{aligned} \tag{7.4.2}
$$

其中

$$
\begin{aligned}
& m_2 = \frac{2\tau + \lambda\tau^2}{4}c, \quad a_2 = \frac{16 + 4\tau^2 - \lambda^2\tau^4}{8\tau - 4\lambda\tau^2} + \frac{2m_2}{h_x^2} + \frac{2m_2}{h_y^2}, \\
& b_2 = \frac{-16 + 4\tau^2 - \lambda^2\tau^4}{8\tau - 4\lambda\tau^2} + \frac{2m_2}{h_x^2} + \frac{2m_2}{h_y^2}, \quad d_2 = \frac{m_2}{c} + \frac{2m_2}{h_x^2} + \frac{2m_2}{h_y^2}.
\end{aligned}
$$

S4 格式:

$$
\begin{aligned}
& a_3 \mathrm{e}^{\frac{\lambda\tau}{2}} u_{j,k}^{n+1} - \frac{m_3}{h_x^2}\mathrm{e}^{\frac{\lambda\tau}{2}}\left(u_{j+1,k}^{n+1} + u_{j-1,k}^{n+1}\right) - \frac{m_3}{h_y^2}\mathrm{e}^{\frac{\lambda\tau}{2}}\left(u_{j,k+1}^{n+1} + u_{j,k-1}^{n+1}\right) \\
& = 2p_{j,k}^n - b_3 u_{j,k}^n + \frac{m_3}{h_x^2}\left(u_{j+1,k}^n + u_{j-1,k}^n\right) + \frac{m_3}{h_y^2}\left(u_{j,k+1}^n + u_{j,k-1}^n\right),
\end{aligned}
$$

$$
\begin{aligned}
&\mathrm{e}^{\frac{\lambda\tau}{2}}p_{j,k}^{n+1}+\frac{4}{4+\lambda\tau}\mathrm{e}^{\frac{\lambda\tau}{2}}\left[d_3u_{j,k}^{n+1}-\frac{m_3}{h_x^2}\left(u_{j+1,k}^{n+1}+u_{j-1,k}^{n+1}\right)\right.\\
&\qquad\left.-\frac{m_3}{h_y^2}\left(u_{j,k+1}^{n+1}+u_{j,k-1}^{n+1}\right)\right]\\
&\quad=\frac{4-\lambda\tau}{4+\lambda\tau}p_{j,k}^{n}-\frac{4}{4+\lambda\tau}\left[d_3u_{j,k}^{n}-\frac{m_3}{h_x^2}\left(u_{j+1,k}^{n}+u_{j-1,k}^{n}\right)\right.\\
&\qquad\left.-\frac{m_3}{h_y^2}\left(u_{j,k+1}^{n}+u_{j,k-1}^{n}\right)\right],
\end{aligned}\tag{7.4.3}
$$

其中

$$
m_3=\frac{\tau}{2}c,\quad a_3=\frac{16+4\tau^2-\lambda^2\tau^2}{8\tau}+\frac{2m_3}{h_x^2}+\frac{2m_3}{h_y^2},
$$

$$
b_3=\frac{-16+4\tau^2-\lambda^2\tau^2-8\lambda\tau}{8\tau}+\frac{2m_3}{h_x^2}+\frac{2m_3}{h_y^2},\quad d_3=\frac{\tau}{2}+\frac{2m_3}{h_x^2}+\frac{2m_3}{h_y^2}.
$$

Hamiltonian 格式:

$$
\begin{aligned}
&a_4u_{j,k}^{n+1}-\frac{m_4}{h_x^2}\left(u_{j+1,k}^{n+1}+u_{j-1,k}^{n+1}\right)-\frac{m_4}{h_y^2}\left(u_{j,k+1}^{n+1}+u_{j,k-1}^{n+1}\right)\\
&\quad=2p_{j,k}^{n}-b_4u_{j,k}^{n}+\frac{m_4}{h_x^2}\left(u_{j+1,k}^{n}+u_{j-1,k}^{n}\right)+\frac{m_4}{h_y^2}\left(u_{j,k+1}^{n}+u_{j,k-1}^{n}\right),\\
&p_{j,k}^{n+1}+\frac{2}{2+\lambda\tau}\left[d_4u_{j,k}^{n+1}-\frac{m_4}{h_x^2}\left(u_{j+1,k}^{n+1}+u_{j-1,k}^{n+1}\right)-\frac{m_4}{h_y^2}\left(u_{j,k+1}^{n+1}+u_{j,k-1}^{n+1}\right)\right]\\
&\quad=\frac{2-\lambda\tau}{2+\lambda\tau}p_{j,k}^{n}-\frac{2}{2+\lambda\tau}\left[d_4u_{j,k}^{n}-\frac{m_4}{h_x^2}\left(u_{j+1,k}^{n}+u_{j-1,k}^{n}\right)\right.\\
&\qquad\left.-\frac{m_4}{h_y^2}\left(u_{j,k+1}^{n}+u_{j,k-1}^{n}\right)\right],
\end{aligned}\tag{7.4.4}
$$

其中

$$
m_4=\frac{\tau}{2}c,\quad a_4=\frac{\tau^2+\lambda\tau+2}{2\tau}+\frac{2m_4}{h_x^2}+\frac{2m_4}{h_y^2},
$$

$$
b_4=\frac{\tau^2-\lambda\tau-2}{2\tau}+\frac{2m_4}{h_x^2}+\frac{2m_4}{h_y^2},\quad d_4=\frac{\tau}{2}+\frac{2m_4}{h_x^2}+\frac{2m_4}{h_y^2}.
$$

取方程的初始值为

$$
\begin{aligned}
&u(0,x,y)=\sin(\omega(x+y)),\\
&u_t(0,x,y)=2\pi f\cos(\omega(x+y))-\frac{\lambda}{2}\sin(\omega(x+y)),
\end{aligned}
$$

且满足

$$c\omega^2 - (2\pi f)^2 + 1 - \frac{\lambda^2}{4} = 0,$$

其中 $\sqrt{c}$ 是波速, f 表示频率. 这个初值问题有精确解

$$u(t, x, y) = \exp\left(-\frac{\lambda t}{2}\right)\sin(\omega(x+y) + 2\pi f t). \tag{7.4.5}$$

数值解的相对误差定义如下:

$$\mathrm{err}(t_n) = \left(\frac{1}{\sum_{j=1}^{J}\sum_{k=1}^{K} u(t_n, x_j, y_k)^2}\sum_{j=1}^{J}\sum_{k=1}^{K}(u_{j,k}^n - u(t_n, x_j, y_k))^2\right)^{\frac{1}{2}},$$

其中 J 和 K 分别表示沿 x 方向和 y 方向的总节点数.

本数值实验的空间区域取为矩形域 $\left[-\frac{\pi}{3}, \frac{\pi}{3}\right] \times \left[-\frac{\pi}{3}, \frac{\pi}{3}\right]$, 时间长为 4π, 空间步长取 $\tau = h_x = h_y = \frac{\pi}{50}$, 波速和频率始终取为 $\sqrt{c} = 200$ 和 $f = 5$, 耗散常数 λ 依次取为 0.1, 0.5, 1.0 和 2.0. 耗散系数不同, 格式展现不同的计算效果.

图 7.4.1—7.4.4 比较了四个格式在不同耗散系数时的误差. Hamiltonian 格式 (7.4.4) 是辛格式和多辛格式, 而耗散系数越小, Birkhoff 系

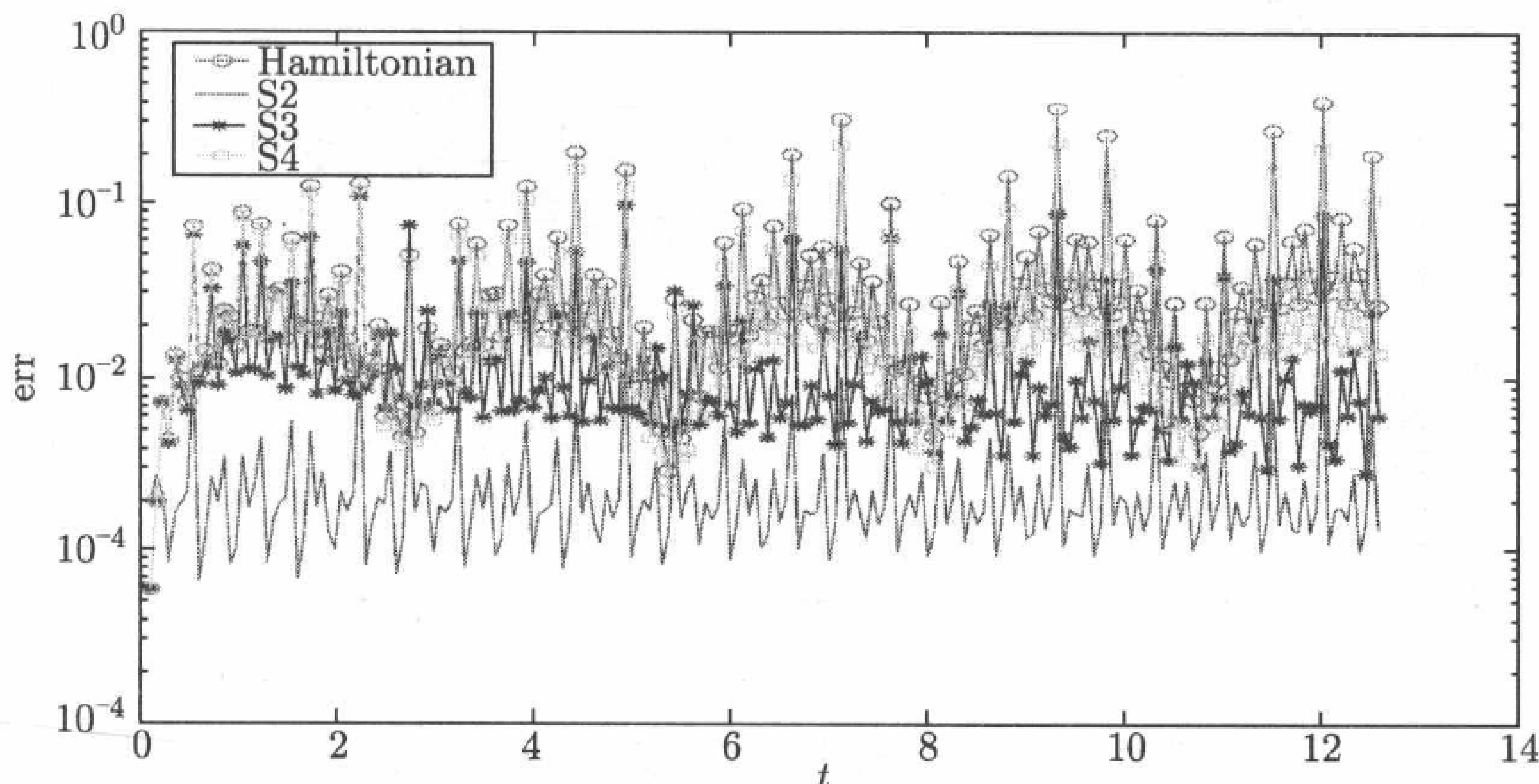


图 7.4.1 $\lambda = 0.1$ 时不同差分格式的误差比较

统越接近一个 Hamilton 系统. 所以, 对于相对较小的 λ, Hamiltonian 格式表现得几乎和所有 Birkhoff 格式一样好; 而对于较大的 λ, 比如 $\lambda = 1.0$ 和 $\lambda = 2.0$, Birkhoff 辛格式尤其是 S2 格式和 S4 格式的优势明显, 而 Hamiltonian 格式始终是最差的格式.

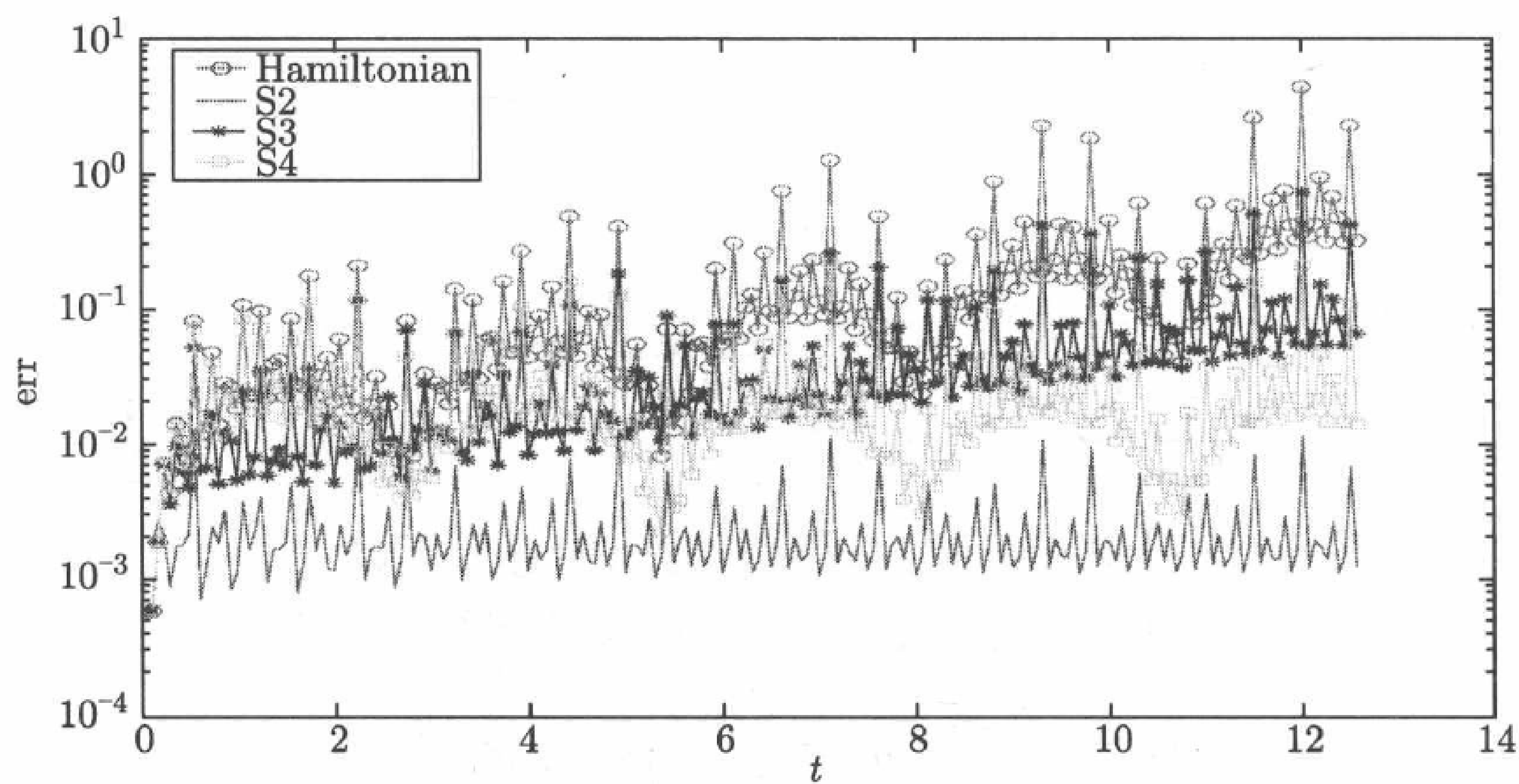


图 7.4.2　$\lambda = 0.5$ 时不同差分格式的误差比较

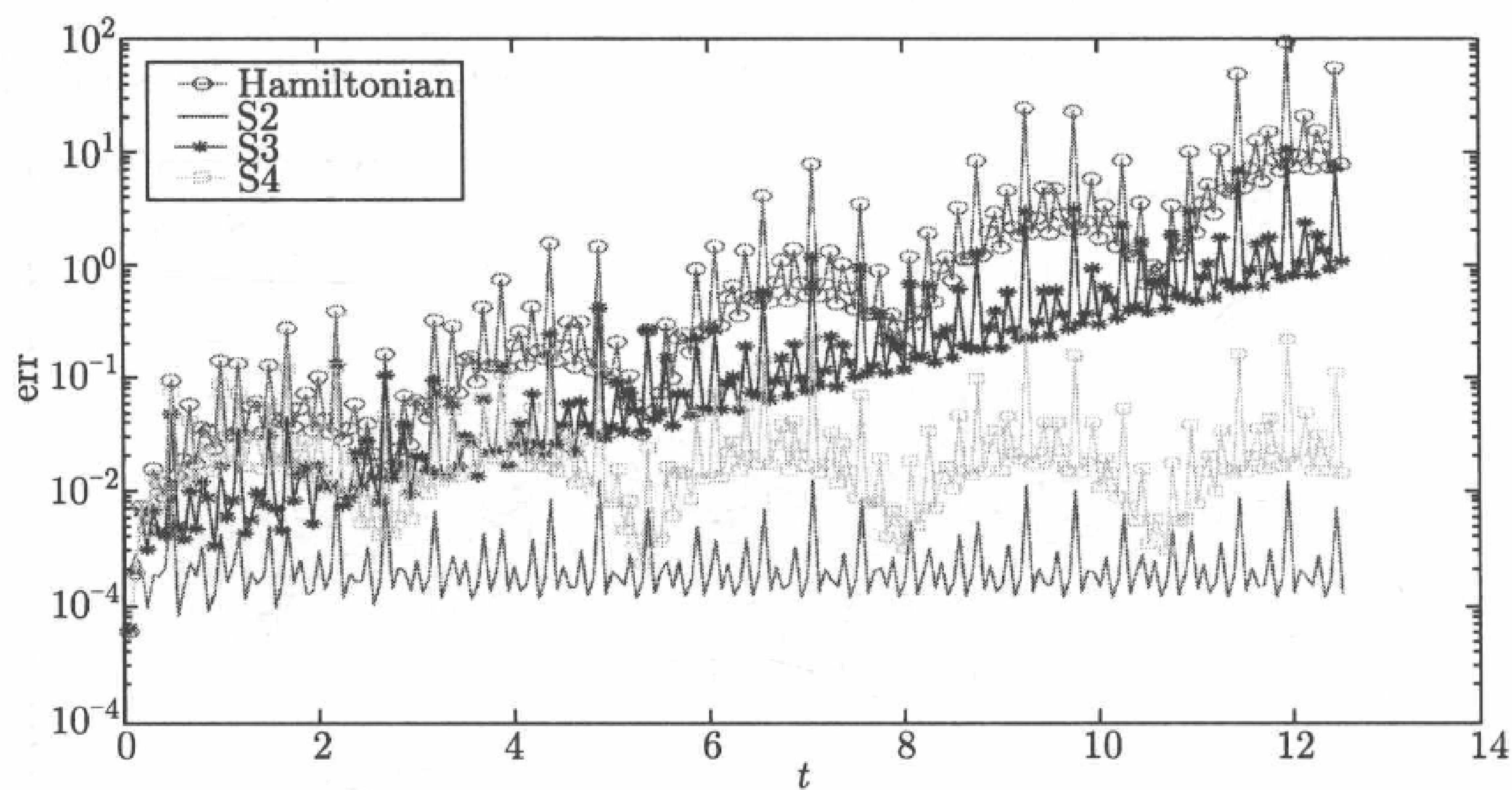


图 7.4.3　$\lambda = 1$ 时不同差分格式的误差比较

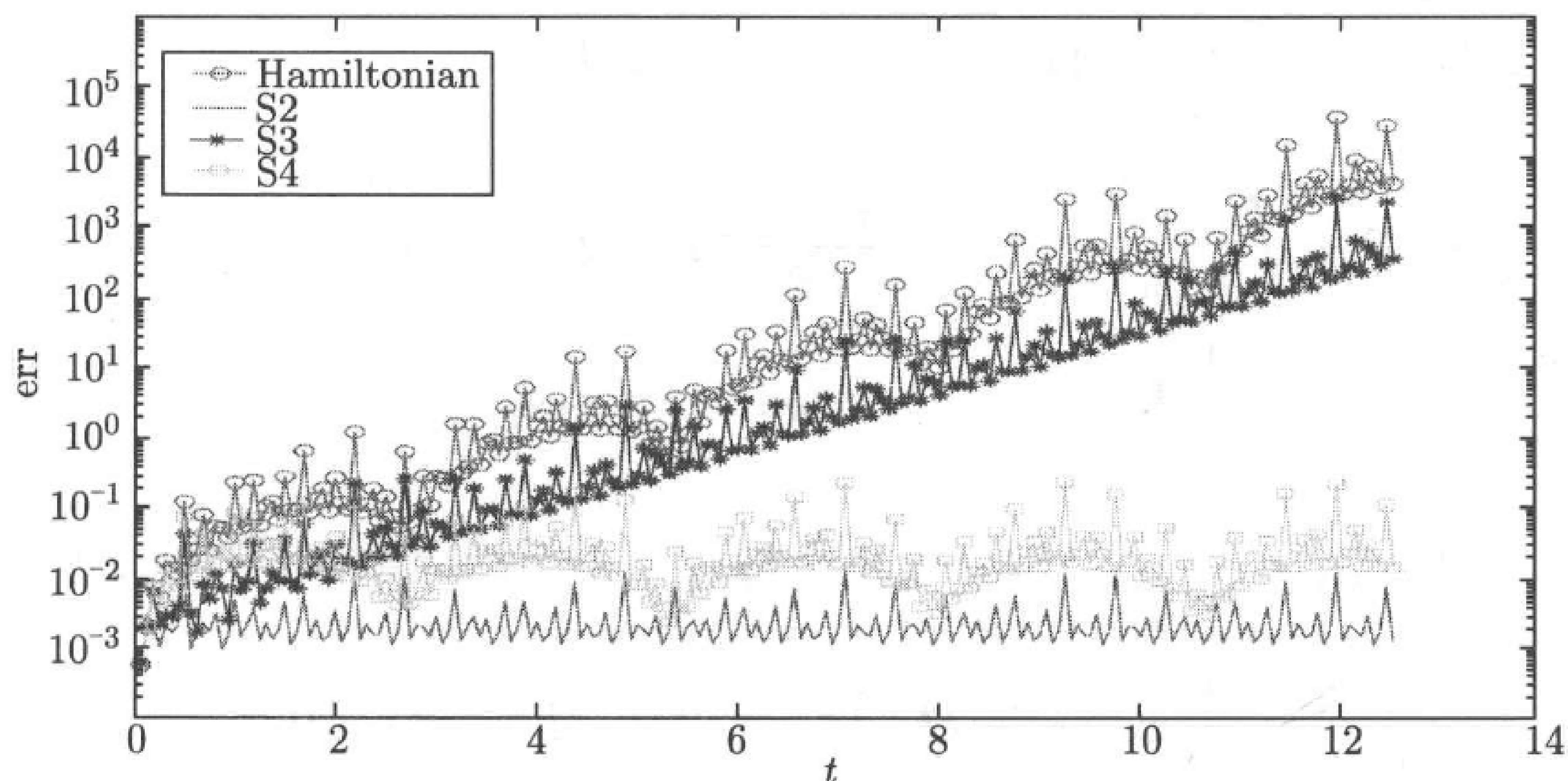


图 7.4.4 $\lambda = 2$ 时不同差分格式的误差比较

7.4.2 TE Maxwell 方程

本小节我们用 Birkhoff 算法模拟带耗散的电磁场方程

$$
\begin{aligned}
&-\frac{\partial \boldsymbol{E}}{\partial t} + \frac{1}{\varepsilon}\nabla \times \boldsymbol{H} - \sigma_1 \boldsymbol{E} = \boldsymbol{0},\\
&\frac{\partial \boldsymbol{H}}{\partial t} + \frac{1}{\mu}\nabla \times \boldsymbol{E} + \sigma_2 \boldsymbol{H} = \boldsymbol{0},
\end{aligned} \tag{7.4.6}
$$

其中 $\boldsymbol{H} = (H_1, H_2, H_3)^{\mathrm{T}}$ 和 $\boldsymbol{E} = (E_1, E_2, E_3)^{\mathrm{T}}$ 分别为磁场和电场强度, 它们都是关于 (x, y, z, t) 的向量函数, μ 为介磁常数, ε 为介电常数, $\sigma_1 = \dfrac{\sigma}{\varepsilon}$, $\sigma_2 = \dfrac{\sigma_m}{\mu}$, 而 σ 和 σ_m 分别为电导率和磁导率.

电磁场方程 (7.4.6) 有无限维 Birkhoff 形式

$$
\boldsymbol{M}(t)\frac{\partial \boldsymbol{Z}}{\partial t} = \delta\mathcal{B}(\boldsymbol{Z}, x, y, z, t) + \frac{\partial \boldsymbol{F}(\boldsymbol{Z}, x, y, z, t)}{\partial t}, \tag{7.4.7}
$$

其中 $\boldsymbol{Z} = (\boldsymbol{H}^{\mathrm{T}}, \boldsymbol{E}^{\mathrm{T}})^{\mathrm{T}}, \boldsymbol{M}(t) = \mathrm{e}^{(\sigma_1+\sigma_2)t}\begin{pmatrix} \boldsymbol{0} & -\boldsymbol{I}_{3\times 3} \\ \boldsymbol{I}_{3\times 3} & \boldsymbol{0} \end{pmatrix}$, 向量函数

$\boldsymbol{F}=\frac{1}{2}\mathrm{e}^{(\sigma_1+\sigma_2)t}(\boldsymbol{E}^{\mathrm{T}},-\boldsymbol{H}^{\mathrm{T}})^{\mathrm{T}}$, Birkhoff 泛函为

$$\begin{aligned}\mathcal{B}=\int_{(x,y,z)}&\mathrm{e}^{(\sigma_1+\sigma_2)t}\left(\frac{\sigma_1-\sigma_2}{2}\boldsymbol{E}\cdot\boldsymbol{H}-\frac{1}{2\varepsilon}\boldsymbol{H}\cdot\nabla\times\boldsymbol{H}\right.\\&\left.-\frac{1}{2\mu}\boldsymbol{E}\cdot\nabla\times\boldsymbol{E}\right)\mathrm{d}x\mathrm{d}y\mathrm{d}z.\end{aligned}\tag{7.4.8}$$

如果 $\sigma_1=\sigma_2=\bar{\sigma}$, 则电磁场方程有一个二次局部能量守恒律:

$$\frac{\partial}{\partial t}\left[\mathrm{e}^{2\bar{\sigma}t}\left(\frac{\varepsilon}{2}|\boldsymbol{E}|^2+\frac{\mu}{2}|\boldsymbol{H}|^2\right)\right]+\mathrm{e}^{2\bar{\sigma}t}\nabla\cdot(\boldsymbol{E}\times\boldsymbol{H})=0.\tag{7.4.9}$$

如果电磁场方程满足周期边界条件, 那么方程有如下整体能量守恒律:

$$\frac{\mathrm{d}\mathcal{H}_1}{\mathrm{d}t}=0,\quad \mathcal{H}_1=\int_{(x,y,z)}\mathrm{e}^{2\bar{\sigma}t}(\varepsilon|\boldsymbol{E}|^2+\mu|\boldsymbol{H}|^2)\mathrm{d}x\mathrm{d}y\mathrm{d}z.\tag{7.4.10}$$

这个守恒律的物理意义是电磁场的 Poynting 能量随时间指数衰减.

本数值实验具体是模拟下面的 TE Maxwell 方程:

$$\begin{aligned}\frac{\partial E_1}{\partial t}&=\frac{1}{\varepsilon}\frac{\partial H_3}{\partial y}-\sigma_1E_1,\\ \frac{\partial E_2}{\partial t}&=-\frac{1}{\varepsilon}\frac{\partial H_3}{\partial x}-\sigma_1E_2,\\ \frac{\partial H_3}{\partial t}&=\frac{1}{\mu}\left(\frac{\partial E_1}{\partial y}-\frac{\partial E_2}{\partial x}\right)-\sigma_2H_3.\end{aligned}\tag{7.4.11}$$

时间方向采用 S4 格式离散, 空间方向采用中点格式离散, 得到格式

$$\begin{aligned}\overline{\partial}_tR_xR_yE_{1i,j}^k&=\frac{1}{\varepsilon}\partial_y\overline{R}_tR_xH_{3i,j}^k-\frac{\sigma_1-\sigma_2}{2}\overline{R}_tR_xR_yE_{1i,j}^k,\\ \overline{\partial}_tR_xR_yE_{2i,j}^k&=-\frac{1}{\varepsilon}\partial_x\overline{R}_tR_yH_{3i,j}^k-\frac{\sigma_1-\sigma_2}{2}\overline{R}_tR_xR_yE_{2i,j}^k,\\ \overline{\partial}_tR_xR_yH_{3i,j}^k&=\frac{1}{\mu}\left(\partial_y\overline{R}_tR_xE_{1i,j}^k-\partial_x\overline{R}_tR_yE_{2i,j}^k\right)\\&\quad-\frac{\sigma_1-\sigma_2}{2}\overline{R}_tR_xR_yH_{3i,j}^k,\end{aligned}\tag{7.4.12}$$

格式中的差分算子为

$$
\begin{aligned}
&\partial_x Q_{i,j}^k = \frac{1}{h_x}\left(Q_{i+1,j}^k - Q_{i,j}^k\right), \quad \partial_y Q_{i,j}^k = \frac{1}{h_y}\left(Q_{i,j+1}^k - Q_{i,j}^k\right),\\
&R_x Q_{i,j}^k = \frac{1}{2}\left(Q_{i,j}^k + Q_{i+1,j}^k\right), \quad R_y Q_{i,j}^k = \frac{1}{2}\left(Q_{i,j}^k + Q_{i,j+1}^k\right), \qquad (7.4.13)\\
&\overline{\partial}_t Q_{i,j}^k = \frac{1}{\tau}\mathrm{e}^{\frac{\sigma_1+\sigma_2}{2}\tau}\left(Q_{i,j}^{k+1} - Q_{i,j}^k\right), \quad \overline{R}_t Q_{i,j}^k = \frac{1}{2}\mathrm{e}^{\frac{\sigma_1+\sigma_2}{2}\tau}\left(Q_{i,j}^{k+1} + Q_{i,j}^k\right).
\end{aligned}
$$

离散的局部能量守恒律为

$$
\begin{aligned}
&\frac{1}{2}\overline{\partial}_t\left[\varepsilon\left(|R_xR_yE_{1i,j}^k|^2 + |R_xR_yE_{2i,j}^k|^2\right) + \mu|R_xR_yH_{3i,j}^k|^2\right]\\
&\quad + \partial_x\left(\overline{R}_tR_yH_{3i,j}^k\overline{R}_tR_yE_{2i,j}^k\right) - \partial_y\left(\overline{R}_tR_xH_{3i,j}^k\overline{R}_tR_xE_{1i,j}^k\right) = 0,
\end{aligned}
\tag{7.4.14}
$$

整体能量守恒律为

$$
\sum_{i,j}\overline{\partial}_t\left[\varepsilon\left(|R_xR_yE_{1i,j}^k|^2 + |R_xR_yE_{2i,j}^k|^2\right) + \mu|R_xR_yH_{3i,j}^k|^2\right] = 0.
\tag{7.4.15}
$$

作为对比, 我们也用指数 FDTD 方法[Ber94] 模拟 TE Maxwell 方程, 其中空间区域为 $[0,1]\times[0,1]$, 时间长为 20, 步长取 $\tau = 0.02$, $h_x = h_y = 0.1$. 图 7.4.5 和图 7.4.6 显示局部能量守恒律 (7.4.9) 和整体能量

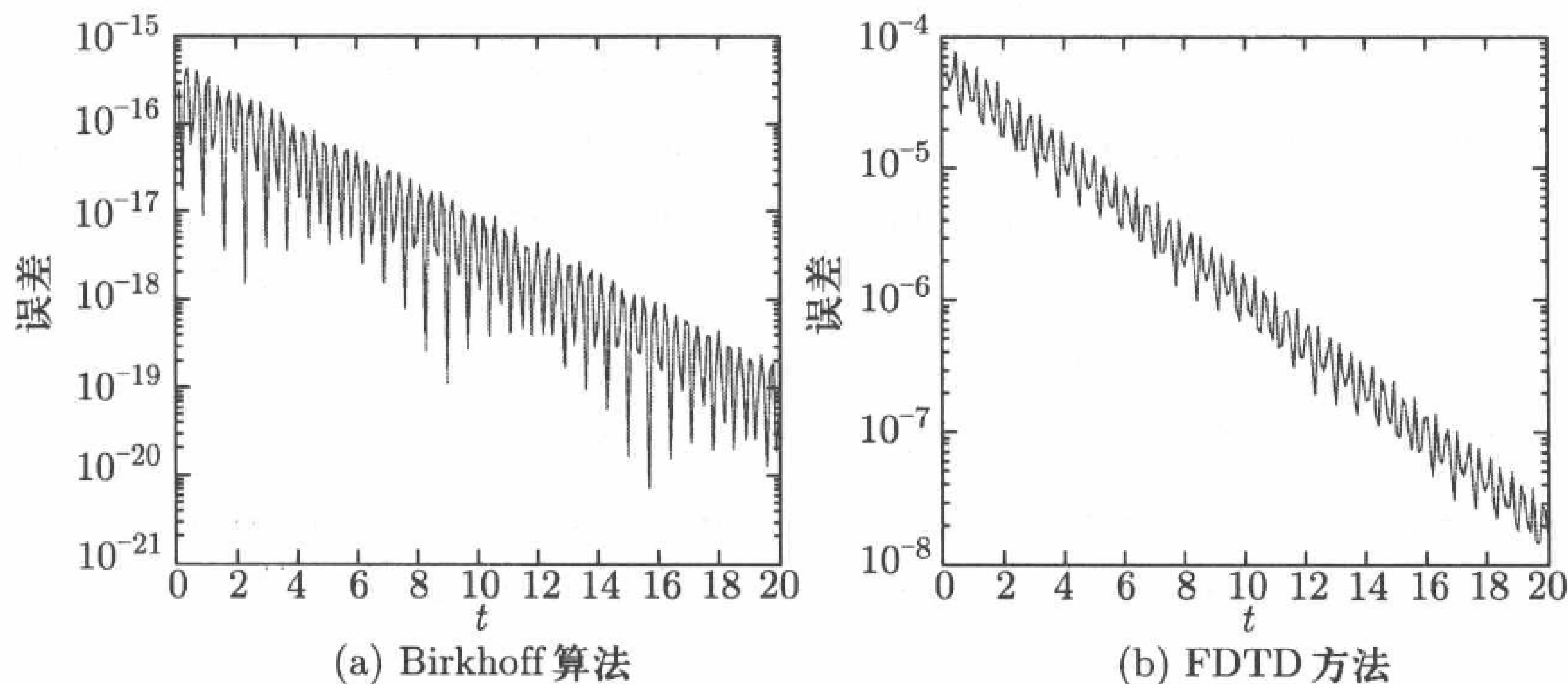
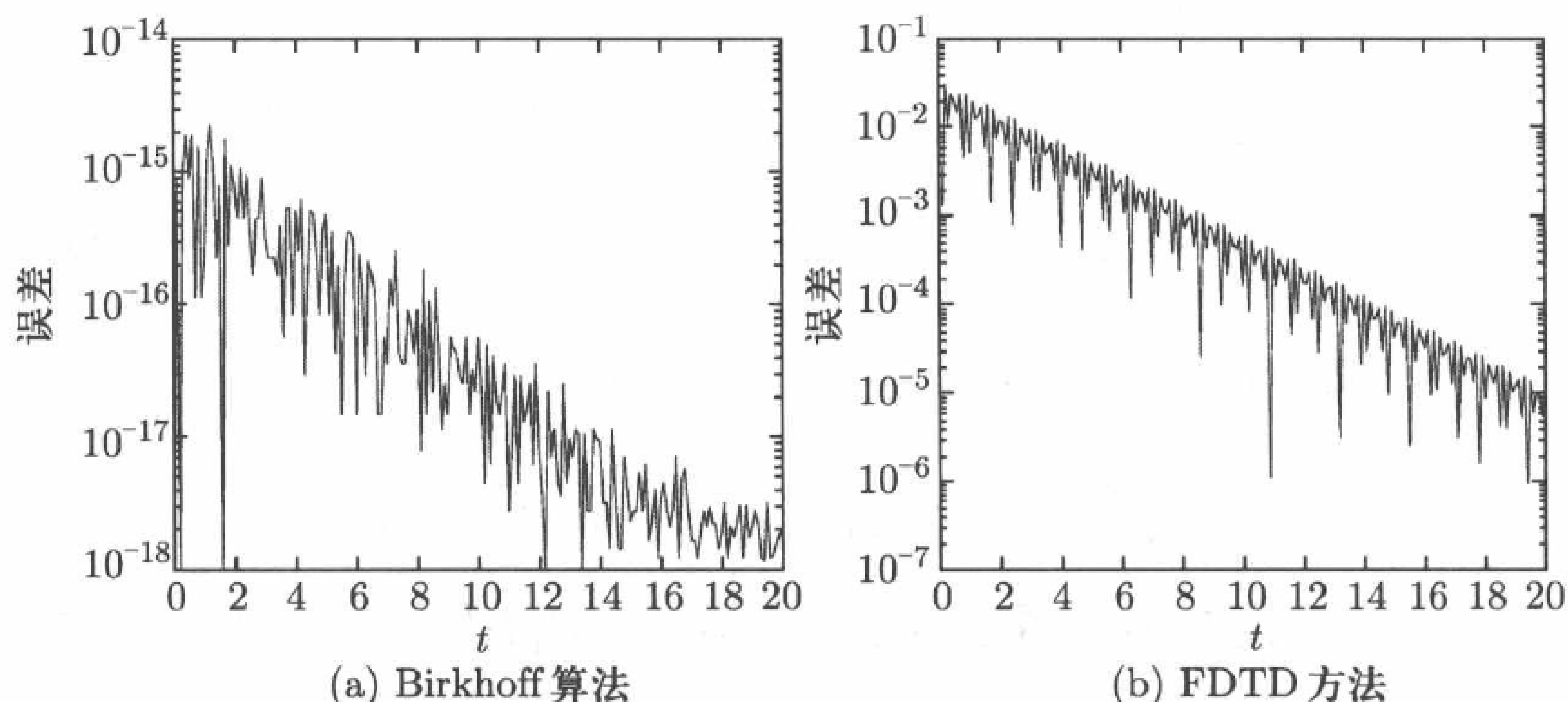


图 7.4.5　局部能量守恒律的误差比较 ($\tau = 0.02$)

守恒律 (7.4.10) 的误差限于机器误差, 所以数值实验结果验证了格式 (7.4.12) 保持局部能量守恒律 (7.4.14) 和整体能量守恒律 (7.4.15).



(a) Birkhoff 算法　　(b) FDTD 方法

图 7.4.6　整体能量守恒律的误差比较 ($\tau = 0.02$)

参 考 文 献

[MW01] Marsden J E, West M. Discrete mechanics and variational integrators. *Acta Numerica*, 2001, 10: 357–514.

[MarRat99] Marsden J E, Ratiu T S. *Introduction to Mechanics and Symmetry*. New York: Springer-Verlag, 1999.

[Arn89] Arnold V I. *Mathematical Methods of Classical Mechanics*. 2nd edition. Berlin: Springer-Verlag, 1989.

[FW91] Feng K, Wang D L. Symplectic difference schemes for Hamiltonian systems in general symplectic structure. *J Comput Math*, 1991, 9: 86–96.

[FQ09] Feng K, Qin M Z. *Symplectic Geometric Algorithms for Hamiltonian Systems*. Zhejiang: Springer-Verlag GmbH, 2009.

[MPS98] Marsden J E, Patrick G P, Shkoller S. Multisymplectic geometry, variational integrators, and nonlinear PDEs. *Commun Math Phys*, 1998, 199: 351–395.

[Bri97] Bridges T J. Multi-symplectic structures and wave propagation. *Math Proc Cam Phil Soc*, 1997, 121: 147–190.

[BR01] Bridges T J, Reich S. Multi-symplectic integrators: numerical schemes for Hamiltonian PDEs that conserve symplecticity. *Phys Lett A*, 2001, 284: 184–193.

[Shang94] Shang Z J. Generating functions for volume-preserving mappings and Hamilton-Jacobi equations for source-free systems. *Science in China (Series A)*, 1994, 37: 1172–1188.

[MS99] Marsden J E, Shkoller S. Multisymplectic geometry, covariant Hamiltonians and water waves. *Math Proc Camb Philos Soc*, 1999, 125: 553–575.

[Tang93] Tang Y F. The necessary condition for a Runge-Kutta scheme to be symplectic for Hamiltonian systems. *Comput Math Appl*, 1993, 26: 13–20.

[McLa94] McLachlan R. Symplectic integration of Hamiltonian wave equations. *Numer Math*, 1994, 66: 465–492.

[HS08] Hong J L, Sun Y J. Generating functions of multi-symplectic RK methods via DW Hamilton-Jacobi equations. *Numer Math*, 2008, 110: 491–519.

[SST02] Saitoh I, Suzuki Y, Takahashi N. Stability of symplectic finite diffrence time domain method. *IEEE Trans Magn*, 2002, 38(6): 665–668.

[Reich02] Reich S. Multi-symplectic Runge-Kutta collocation methods for Hamiltonian wave equations. *J Comput Phys*, 2002, 157: 473–499.

[Olver04] Olver P J, West M, Wulff C. Approximate momentum conservation for spatial semidiscretizations of semilinear wave equations. *Numer Math*, 2004, 97: 493–535.

[Cano06] Cano B. Conserved quantities of some Hamiltonian wave equations after full discretization. *Numer Math*, 2006, 103: 197–223.

[San83b] Santilli R M. *Foundations of Theoretical Mechanics II*. 2nd edition. New York: Springer-Verlag, 1983.

[HLW02] Hairer E, Lubich C, Wanner G. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. Berlin: Springer, 2002.

[Ber94] Berenger J P. A perfectly matched layer for the absorption of electromagnetic waves. *J Comput Phys*, 1994, 114: 185–200.

[SL14] Su Hongling, Li Shengtai. Structure-preserving numerical methods for infinite-dimensional Birkhoffian systems. *J Comput Sci*, 2014, DOI 10.1007/s10915-014-9958-2.

第八章　电磁场方程的多辛结构

电磁场理论在许多现代科技中有很重要的应用. Maxwell 方程作为对电磁场活动的完整描述, 体现着不可替代的作用, 对于它的研究从来就没有冷落过. 关于电磁场方程的数值求解方法, 最常用的是 Yee[Yee66] 于 1966 年提出的时域有限差分方法 (FDTD 方法). 文献 [Yee66] 以 Yee 元胞 (cell) 为空间电磁场离散单元, 将 Maxwell 方程转化为差分方程, 给出了处在一个四边都是绝缘体的长方形中的电磁波的传播和反射的数值模拟. FDTD 方法自提出之后在工程电磁学等领域备受重视, 获得了广泛的应用和发展, 可以说后来很多关于电磁场方程的数值计算研究工作, 都是针对不同的情况对 Yee 的数值逼近的改进. 早期的有关电磁场 FDTD 方法的文献都可在文献 [Taflove05] 中找到. 另外, 也有很多学者研究电磁流的场论和几何背景, 以掌握电磁场的内在特性, 从而找到能够保持它的某些特性的数值逼近, 即保结构算法.

Maxwell 方程的保结构算法主要有辛算法和多辛算法, 这两种算法都是基于方程的辛形式和多辛形式离散而来的. 我们在文献 [SQ03, SQ07] 中首次详细讨论了电磁场方程的多辛结构, 基于这些多辛结构的电磁场方程, 辛和多辛算法散见于近年的一些文献 [SQ07, KH07, SunT11, CWS13], 这里我们不一一列举.

我们在这一章主要讨论不同物理环境中的 Maxwell 方程的对称性和多辛形式, 使读者能清晰理解和掌握 Maxwell 方程的多辛结构, 并能基于这些结构构造保结构算法.

§8.1　电磁场方程的一阶 Lagrange 形式

在工程和数学物理中, 我们需要解决许多描述集方程. **描述集**是指用方程的变化和方程的构造以及初边值条件等来描述方程. 而被描述的方程往往并非微分方程的形式, 有时这就给数学工具的使用带来

麻烦. 根据某个泛函, 可以提供描述集的替代公式, 这一过程称为变分过程. 变分过程往往产生多个 Euler-Lagrange 方程和不同的初边值条件, 如果它们和描述集匹配, 则寻找泛函问题的不动点提供了所研究的物理问题的替代描述. 给出一个微分方程, 寻找它对应的泛函就是变分反问题. 求解变分反问题时, 需要先判断给定的微分方程是否是某个泛函极值问题的 Euler-Lagrange 方程, 如果是, 然后给出它对应的泛函.

判断一个微分方程是否是 Euler-Lagrange 方程, 就是看它对应的导数算子是否是势算子. 我们在第三章给出了 Euler-Lagrange 方程的判别准则: 如果一个微分方程满足对称条件, 则它是 Euler-lagrange 方程. 本章我们引用第三章的方法来讨论 Maxwell 方程的对称性.

考虑一个映射 $\boldsymbol{u}:\mathbf{R}^{1,3}\to\mathbf{R}^6$ 上的 k 阶泛函. 具体地就 Maxwell 方程来说, $\boldsymbol{u}=(H_1,H_2,H_3,E_1,E_2,E_3)$. 这里 $\mathbf{R}^{1,3}$ 是一个 $1+3$ 维的时–空区域, 其局部坐标为 (x_1,x_2,x_3,x_4), 其中 $x_1=t$. $\mathbf{R}^6$ 是目标空间, 用局部坐标表示为 $\boldsymbol{u}=(u^1,\cdots,u^6)$, 把空间 $T_{\boldsymbol{u}}\mathbf{R}^6\times T^*_{\boldsymbol{x}}\mathbf{R}^{1,3}$ 记为 $\mathfrak{R}$. 一个 Lagrange 泛函通过 Lagrange 因子 L 来定义. 假设 Lagrange 因子 L 是定义在空间 $\mathbf{R}^{1,3}\times\mathbf{R}^6\times\mathfrak{R}$ 上的光滑函数, 也可以说 L 是定义在空间 $\mathbf{R}^{1,3}\times\mathbf{R}^6$ 的切丛上的光滑函数, 切丛的纤维 $(\boldsymbol{x},\boldsymbol{u})$ 等价于 $\mathfrak{R}$. 所以 Lagrange 泛函就可以定义为如下形式:

$$\mathcal{L}=\int_{\mathbf{R}^{1,3}}L(x_j,u^i(\boldsymbol{x}),u^i_\nu(\boldsymbol{x}))\omega, \tag{8.1.1}$$

此处 ω 表示 $\mathbf{R}^{1,3}$ 上的 4-形式, $u^i_\nu=\dfrac{\partial u^i}{\partial x_{j_1}\cdots\partial x_{j_\nu}}$ $(\nu=1,\cdots,k)$. 我们定义共轭的 Legendre 变换

$$\pi^\nu_i=\frac{\partial}{\partial u^i_\nu}L(x_j,u^i(\boldsymbol{x}),u^i_\nu(\boldsymbol{x})). \tag{8.1.2}$$

由这个变换可导出定义在多辛流形 $\mathbf{R}^{1,3}\times\mathbf{R}^6\times\mathfrak{R}$ 上的 Hamilton 函数 $\mathcal{H}$:

$$\mathcal{H}(x_j,u^i(\boldsymbol{x}),\pi^\nu_i)=\pi^\nu_i u^i_\nu-L(x_j,u^i(\boldsymbol{x}),u^i_\nu(\boldsymbol{x})). \tag{8.1.3}$$

现在具体地考虑非均匀、各向同性、无损耗的介质中的 Maxwell 方程

$$
\begin{aligned}
&\mu\frac{\partial \boldsymbol{H}}{\partial t}+\nabla\times\boldsymbol{E}=-\boldsymbol{K},\\
&-\nabla\times\boldsymbol{H}+\varepsilon\frac{\partial \boldsymbol{E}}{\partial t}=-\boldsymbol{J},
\end{aligned}
\tag{8.1.4}
$$

其中 $\boldsymbol{E}=(E_1,E_2,E_3)^{\mathrm{T}}$ 是电场强度, $\boldsymbol{H}=(H_1,H_2,H_3)^{\mathrm{T}}$ 是磁场强度, ε 是介电常数, μ 是介磁常数 (在这里我们认为 ε 和 μ 都是关于时间和空间变量的函数), $\boldsymbol{J}$ 是外加电流密度, $\boldsymbol{K}$ 是外加磁流密度.

令 $\boldsymbol{Z}=\boldsymbol{Z}(x,y,z,t)$ 是场向量, 它包括电场强度 $\boldsymbol{E}$ 和磁场强度 $\boldsymbol{H}$ 两个分量, 也就是说 $\boldsymbol{Z}$ 是一个六维向量, 写成

$$
\boldsymbol{Z}=(H_1,H_2,H_3,E_1,E_2,E_3)^{\mathrm{T}},
\tag{8.1.5}
$$

也可将它简单地记为 $\boldsymbol{Z}=(\boldsymbol{H}^{\mathrm{T}},\boldsymbol{E}^{\mathrm{T}})^{\mathrm{T}}$. 再令 $\boldsymbol{F}=\boldsymbol{F}(x,y,z,t)$ 是由外加电流 $\boldsymbol{J}$ 和外加磁流 $\boldsymbol{K}$ 组成的源向量 $(\boldsymbol{J}^{\mathrm{T}},-\boldsymbol{K}^{\mathrm{T}})^{\mathrm{T}}$, 则方程组 (8.1.4) 可以写成下面的形式:

$$
\boldsymbol{G}\cdot\boldsymbol{Z}=\boldsymbol{F},
\tag{8.1.6}
$$

这里 $\boldsymbol{G}$ 是一个算子矩阵:

$$
\boldsymbol{G}=\begin{pmatrix}\nabla\times & -\varepsilon\boldsymbol{\partial}_t\\ \mu\boldsymbol{\partial}_t & \nabla\times\end{pmatrix},
\tag{8.1.7}
$$

其中 $\boldsymbol{G}$ 的每一个元素都是一个三阶算子矩阵, 其中

$$
\nabla\times=\begin{pmatrix}0 & -\dfrac{\partial}{\partial z} & \dfrac{\partial}{\partial y}\\ \dfrac{\partial}{\partial z} & 0 & -\dfrac{\partial}{\partial x}\\ -\dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial x} & 0\end{pmatrix},\quad
\boldsymbol{\partial}_t=\begin{pmatrix}\partial_t & 0 & 0\\ 0 & \partial_t & 0\\ 0 & 0 & \partial_t\end{pmatrix}.
\tag{8.1.8}
$$

显然, 矩阵 $\boldsymbol{G}$ 的每一个元素都是一个线性算子, 我们称这种矩阵为**线性**的. 我们在第三章定义了算子矩阵的微分. 由定义知道, 对于线性算

子矩阵 $\boldsymbol{G}$, 它的微分算子矩阵 $\boldsymbol{G}'$ 一定等于它本身. 由第三章的共轭算子定义知道, 矩阵 $\boldsymbol{G}'$ 的共轭矩阵是

$$\boldsymbol{G}'^{*} = \begin{pmatrix} (\nabla\times)^{\mathrm{T}} & \varepsilon\boldsymbol{\partial}_t + \boldsymbol{\varepsilon}_t \\ -\mu\boldsymbol{\partial}_t - \boldsymbol{\mu}_t & (\nabla\times)^{\mathrm{T}} \end{pmatrix}, \tag{8.1.9}$$

其中 $\boldsymbol{\varepsilon}_t = \varepsilon_t \boldsymbol{I}, \boldsymbol{\mu}_t = \mu_t \boldsymbol{I}$.

如果方程 (8.1.4) 改写成 (8.1.6) 式的形式后满足条件 $\boldsymbol{G}'^{\mathrm{T}} = \boldsymbol{G}'^{*}$, 则称方程 (8.1.4) 的 (8.1.6) 形式是对称的. 此时必须有 $\varepsilon = \mu$, 并且两个系数不能依赖时间变量 t. 真空中的 Maxwell 方程显然满足这个条件, 所以说它的 (8.1.6) 形式是对称的.

因为 (8.1.6) 形式是对称的, 从而对应一个 Lagrange 泛函

$$\begin{aligned} \mathcal{L} &= \int_{\mathbf{R}^{1,3}} L(\boldsymbol{H}, \boldsymbol{E}, \boldsymbol{H}_t, \boldsymbol{E}_t)\omega, \\ L &= \frac{1}{2}\langle \boldsymbol{H}, \nabla\times\boldsymbol{H}\rangle + \frac{1}{2}\langle \boldsymbol{E}, \nabla\times\boldsymbol{E}\rangle \\ &\quad -\mu\langle \boldsymbol{H}, \boldsymbol{E}_t\rangle - \langle \boldsymbol{H}, \boldsymbol{J}\rangle + \langle \boldsymbol{E}, \boldsymbol{K}\rangle. \end{aligned} \tag{8.1.10}$$

我们知道, 对于同一形式的 (8.1.6) 式, Lagrange 函数并不唯一, 还有其他类似的 Lagrange 函数, 比如

$$\begin{aligned} L_1 = {} & \frac{1}{2}\langle \boldsymbol{H}, \nabla\times\boldsymbol{H}\rangle + \frac{1}{2}\langle \boldsymbol{E}, \nabla\times\boldsymbol{E}\rangle - \frac{1}{2}\mu\langle \boldsymbol{H}, \boldsymbol{E}_t\rangle \\ & + \frac{1}{2}\varepsilon\langle \boldsymbol{H}_t, \boldsymbol{E}\rangle - \langle \boldsymbol{H}, \boldsymbol{J}\rangle + \langle \boldsymbol{E}, \boldsymbol{K}\rangle, \end{aligned} \tag{8.1.11}$$

命题 8.1.1　当 $\mu = \mu(x,y,z) = \varepsilon(x,y,z) = \varepsilon$ 时, 以形式 (8.1.4) 表示的 Maxwell 方程 (8.1.4) 是 Lagrange 函数 (8.1.10) 和 (8.1.11) 对应的变分问题的 Euler-Lagrange 方程.

命题由第三章的定理 3.1.1 即可证得.

现在考虑方程 (8.1.4) 的另外一种形式:

$$\begin{aligned} & \boldsymbol{G}_1 \cdot \boldsymbol{Z} = \boldsymbol{F}_1, \\ & \boldsymbol{G}_1 = \begin{pmatrix} \frac{1}{\varepsilon}\nabla\times & -\boldsymbol{\partial}_t \\ \boldsymbol{\partial}_t & \frac{1}{\mu}\nabla\times \end{pmatrix}, \quad \boldsymbol{F}_1 = \left(\frac{1}{\varepsilon}\boldsymbol{J}^{\mathrm{T}}, -\frac{1}{\mu}\boldsymbol{K}^{\mathrm{T}}\right)^{\mathrm{T}}. \end{aligned} \tag{8.1.12}$$

因为矩阵 $\boldsymbol{G}_1$ 仍然是线性算子, 所以它的微分算子矩阵等于它本身, 它的共轭算子矩阵是

$$\boldsymbol{G}_1'^* = \begin{pmatrix} \boldsymbol{A} + \dfrac{1}{\varepsilon}(\nabla\times)^{\mathrm{T}} & \boldsymbol{\partial}_t \\ -\boldsymbol{\partial}_t & \boldsymbol{B} + \dfrac{1}{\mu}(\nabla\times)^{\mathrm{T}} \end{pmatrix}, \tag{8.1.13}$$

这里 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是三阶矩阵:

$$\boldsymbol{A} = \begin{pmatrix} 0 & \left(\dfrac{1}{\varepsilon}\right)_z & -\left(\dfrac{1}{\varepsilon}\right)_y \\ -\left(\dfrac{1}{\varepsilon}\right)_z & 0 & -\left(\dfrac{1}{\varepsilon}\right)_x \\ \left(\dfrac{1}{\varepsilon}\right)_y & -\left(\dfrac{1}{\varepsilon}\right)_x & 0 \end{pmatrix},$$
$$\boldsymbol{B} = \begin{pmatrix} 0 & \left(\dfrac{1}{\mu}\right)_z & -\left(\dfrac{1}{\mu}\right)_y \\ -\left(\dfrac{1}{\mu}\right)_z & 0 & -\left(\dfrac{1}{\mu}\right)_x \\ \left(\dfrac{1}{\mu}\right)_y & -\left(\dfrac{1}{\mu}\right)_x & 0 \end{pmatrix}. \tag{8.1.14}$$

因此若要满足对称条件 $\boldsymbol{G}_1'^{\mathrm{T}} = \boldsymbol{G}_1'^*$, 参数 ε 和 μ 必须不依赖于时间变量 t, 并且不能等于零. 这样我们得到系统 (8.1.4) 的另一个 Lagrange 函数

$$\begin{aligned} L_2 = {} & \frac{1}{2\varepsilon}\langle \boldsymbol{H}, \nabla\times\boldsymbol{H}\rangle + \frac{1}{2\mu}\langle \boldsymbol{E}, \nabla\times\boldsymbol{E}\rangle \\ & -\langle \boldsymbol{H}, \boldsymbol{E}_t\rangle - \frac{1}{\varepsilon}\langle \boldsymbol{H}, \boldsymbol{J}\rangle + \frac{1}{\mu}\langle \boldsymbol{E}, \boldsymbol{K}\rangle. \end{aligned} \tag{8.1.15}$$

命题 8.1.2 当 $\mu = \mu(x,y,z) \neq 0$, 且 $\varepsilon = \varepsilon(x,y,z) \neq 0$ 时, 以形式 (8.1.12) 表示的 Maxwell 方程 (8.1.4) 是 Lagrange 函数 (8.1.15) 对应的变分问题的 Euler-Lagrange 方程.

命题同样可由第三章的定理 3.1.1 证得.

从形式上来看, Lagrange 函数 L_2 和 L 除了参数 ε 和 μ 的位置不同之外, 差别不大. 但是事实上, 因为对介电常数和磁导率的约束条件

不同, 使得 L_2 和 L 是针对不同的系统得来的, 由它们定义的泛函极值问题也是对不同的物理现象的替代描述.

以上给出的 Lagrange 函数都是一阶的, 从而泛函极值问题对应的 Euler-Lagrange 方程也只有一阶, 因此用共轭 Legendre 变换求出它的广义共轭动量是平凡的, 比如我们关于 L_2 求共轭动量, 得到

$$\begin{aligned} &\frac{\partial L_2}{\partial \nabla \times \boldsymbol{H}} = \frac{1}{2\varepsilon}\boldsymbol{H}, \\ &\frac{\partial L_2}{\partial \nabla \times \boldsymbol{E}} = \frac{1}{2\mu}\boldsymbol{E}, \\ &\frac{\partial L_2}{\partial \boldsymbol{E}_t} = -\boldsymbol{H}. \end{aligned} \tag{8.1.16}$$

显然, 求出的共轭动量只是在原来的变量上乘以一个参数, 这是没有意义的. 事实上, 共轭 Legendre 变换是通过引进广义共轭动量, 使得原来的 Euler-Lagrange 方程得以降低阶数, 变成一个一阶的 Hamilton 方程. 因此, 对于一个本身就是一阶方程的系统来说, 假如它是 Euler-Lagrange 方程, 也就是说它是对称的, 那么只有两种结果, 要么它已经是一个 Hamilton 方程, 要么它就是非自治的 Birkhoff 方程. 方程 (8.1.4) 就是 Hamilton 方程.

§8.2　电磁场方程的多辛 Hamilton 形式

从上一节的讨论知道, 当 ε 和 μ 满足一定条件时, Maxwell 方程是对称的, 又因为它是一阶方程, 所以它本身就是一个 Hamilton 形式. 但是, 当 ε 和 μ 不满足命题 8.1.1 和命题 8.1.2 的条件时, Maxwell 方程 (8.1.4) 可能是不对称的. 对于不对称的一阶方程, 我们采取引进势函数的办法, 使它变成一个对称的高阶方程. 引进势函数使方程 (8.1.4) 变成一个二阶对称方程, 便可以找到二阶 Lagrange 函数, 再通过 Legendre 变换对方程进行降阶, 可得到和原方程等价的一阶 Hamilton 形式.

引进两个向量势函数 $\boldsymbol{U} = (U_1, U_2, U_3)$ 和 $\boldsymbol{V} = (V_1, V_2, V_3)^{\mathrm{T}}$, 使得 $\boldsymbol{U}_t = \boldsymbol{E}$, $\boldsymbol{V}_t = \boldsymbol{H}$. 令 $\boldsymbol{Z} = (V_1, V_2, V_3, U_1, U_2, U_3)^{\mathrm{T}}$. 方程 (8.1.4) 可以

改写为如下形式:

$$G_2 \cdot Z = F_2, \tag{8.2.1}$$

其中

$$G_2 = \begin{pmatrix} \mu\partial_t^2 & \nabla \times \partial_t \\ -\nabla \times \partial_t & \varepsilon\partial_t^2 \end{pmatrix}, \quad F_2 = -(K^{\mathrm{T}}, J^{\mathrm{T}})^{\mathrm{T}}.$$

因为矩阵 G_2 也是线性的, 所以有 $G_2' = G_2$. 如果 ε 和 μ 不依赖时间变量 t, 那么很显然 G_2' 是自共轭的. 这时对应的 Lagrange 函数为

$$\begin{aligned} L(V, U) &= \int_X \int_0^1 < [V, U], G_2(\lambda Z) > \mathrm{d}\lambda \mathrm{d}V \\ &= \int_X \left(\frac{1}{2}\mu\langle V, V_{tt}\rangle + \frac{1}{2}\langle V, \nabla \times U_t\rangle + \frac{1}{2}\varepsilon\langle U, U_{tt}\rangle \right. \\ &\quad \left. -\frac{1}{2}\langle U, \nabla \times V_t\rangle + U^{\mathrm{T}}J + V^{\mathrm{T}}K \right) \mathrm{d}V \\ &= \int_X \left(-\frac{1}{2}\mu\langle V_t, V_t\rangle - \frac{1}{2}\langle V_t, \nabla \times U\rangle - \frac{1}{2}\varepsilon\langle U_t, U_t\rangle \right. \\ &\quad \left. +\frac{1}{2}\langle U_t, \nabla \times V\rangle + U^{\mathrm{T}}J + V^{\mathrm{T}}K + \operatorname{div} P \right) \mathrm{d}V, \end{aligned} \tag{8.2.2}$$

其中 X 是 $\mathbf{R}^{1,3}$ 的一个开子集, $\operatorname{div} P$ 是任意散度函数.

命题 8.2.1 当 ε 和 μ 不依赖于时间 t 时, 方程 (8.2.1) 是变分问题 (8.2.2) 的 Euler-Lagrange 方程.

命题由第三章的定理 3.1.1 即可证得.

对一个给定的 Euler-Lagrange 方程, 它的 Lagrange 函数在相差一个散度项 $\operatorname{div} P$ 的情况下是唯一的, 所以我们取其中一个 Lagrange 函数:

$$\begin{aligned} \widetilde{L} &= \frac{1}{2}\mu\langle V_t, V_t\rangle + \frac{1}{2}\langle V_t, \nabla \times U\rangle + \frac{1}{2}\varepsilon\langle U_t, U_t\rangle \\ &\quad -\frac{1}{2}\langle U_t, \nabla \times V\rangle - U^{\mathrm{T}}J - V^{\mathrm{T}}K. \end{aligned} \tag{8.2.3}$$

引进 $\widetilde{L}$ 的广义共轭动量

$$\begin{aligned}
&\boldsymbol{P}=\frac{\partial\widetilde{L}}{\partial\boldsymbol{V}_t}=\mu\boldsymbol{V}_t+\frac{1}{2}\nabla\times\boldsymbol{U},\quad \frac{\partial\widetilde{L}}{\partial\nabla\times\boldsymbol{V}}=-\frac{1}{2}\boldsymbol{U}_t=-\frac{1}{2}\boldsymbol{E},\\
&\boldsymbol{Q}=\frac{\partial\widetilde{L}}{\partial\boldsymbol{U}_t}=\varepsilon\boldsymbol{U}_t-\frac{1}{2}\nabla\times\boldsymbol{V},\quad \frac{\partial\widetilde{L}}{\partial\nabla\times\boldsymbol{U}}=\frac{1}{2}\boldsymbol{V}_t=\frac{1}{2}\boldsymbol{H}.
\end{aligned}\tag{8.2.4}$$

用 $\overline{\boldsymbol{Z}}$ 表示 18 维的向量 $(\boldsymbol{H}^{\mathrm{T}},\boldsymbol{E}^{\mathrm{T}},\boldsymbol{V}^{\mathrm{T}},\boldsymbol{U}^{\mathrm{T}},\boldsymbol{P}^{\mathrm{T}},\boldsymbol{Q}^{\mathrm{T}})^{\mathrm{T}}$, 则通过下面的计算可以得到一个协变 Hamilton 量 $S(\overline{\boldsymbol{Z}})$:

$$\begin{aligned}
S(\overline{\boldsymbol{Z}})=&\left\langle\frac{\partial\widetilde{L}}{\partial\boldsymbol{V}_t},\boldsymbol{V}_t\right\rangle+\left\langle\frac{\partial\widetilde{L}}{\partial\boldsymbol{U}_t},\boldsymbol{U}_t\right\rangle\\
&+\left\langle\frac{\partial\widetilde{L}}{\partial\nabla\times\boldsymbol{V}},\nabla\times\boldsymbol{V}\right\rangle+\left\langle\frac{\partial\widetilde{L}}{\partial\nabla\times\boldsymbol{U}},\nabla\times\boldsymbol{U}\right\rangle-\widetilde{L}\\
=&\langle\boldsymbol{P},\boldsymbol{V}_t\rangle+\langle\boldsymbol{Q},\boldsymbol{U}_t\rangle+\left\langle\frac{\partial\widetilde{L}}{\partial\nabla\times\boldsymbol{V}},\nabla\times\boldsymbol{V}\right\rangle+\left\langle\frac{\partial\widetilde{L}}{\partial\nabla\times\boldsymbol{U}},\nabla\times\boldsymbol{U}\right\rangle-\widetilde{L}\\
=&\langle\boldsymbol{P},\boldsymbol{H}\rangle+\langle\boldsymbol{Q},\boldsymbol{E}\rangle-\frac{1}{2}\mu\langle\boldsymbol{H},\boldsymbol{H}\rangle-\frac{1}{2}\varepsilon\langle\boldsymbol{E},\boldsymbol{E}\rangle+\boldsymbol{U}^{\mathrm{T}}\boldsymbol{J}+\boldsymbol{V}^{\mathrm{T}}\boldsymbol{K}.
\end{aligned}\tag{8.2.5}$$

根据 (8.2.4) 式和 (8.2.5) 式, 我们得到了 Maxwell 方程 (8.1.4) 的一个多辛 Hamilton 形式

$$\begin{aligned}
&\frac{1}{2}\nabla\times\boldsymbol{U}=\boldsymbol{P}-\mu\boldsymbol{H},\quad -\frac{1}{2}\nabla\times\boldsymbol{V}=\boldsymbol{Q}-\varepsilon\boldsymbol{E},\\
&-\boldsymbol{P}_t-\frac{1}{2}\nabla\times\boldsymbol{E}=\boldsymbol{K},\quad -\boldsymbol{Q}_t+\frac{1}{2}\nabla\times\boldsymbol{H}=\boldsymbol{J},\\
&\boldsymbol{V}_t=\boldsymbol{H},\quad \boldsymbol{U}_t=\boldsymbol{E}.
\end{aligned}\tag{8.2.6}$$

把上面的方程写成如下矩阵形式:

$$\boldsymbol{\mathcal{M}}\overline{\boldsymbol{Z}}_t+\boldsymbol{\mathcal{K}}\nabla\times\overline{\boldsymbol{Z}}=\nabla_{\overline{\boldsymbol{Z}}}S(\overline{\boldsymbol{Z}}),\tag{8.2.7}$$

其中 $\nabla\times\overline{\boldsymbol{Z}}$ 表示对 $\overline{\boldsymbol{Z}}$ 的各个变量作旋度, 即

$$(\nabla\times\boldsymbol{H},\nabla\times\boldsymbol{E},\nabla\times\boldsymbol{V},\nabla\times\boldsymbol{U},\nabla\times\boldsymbol{P},\nabla\times\boldsymbol{Q})^{\mathrm{T}},$$

而 $\mathcal{K}\nabla\times\overline{\boldsymbol{Z}}$ 由三个部分相加:

$$\mathcal{K}\nabla\times\overline{\boldsymbol{Z}}=\mathcal{K}_1\overline{\boldsymbol{Z}}_x+\mathcal{K}_2\overline{\boldsymbol{Z}}_y+\mathcal{K}_3\overline{\boldsymbol{Z}}_z, \tag{8.2.8}$$

这里矩阵 $\mathcal{M}$ 和 $\mathcal{K}_i$ 分别是

$$\mathcal{M}=\begin{pmatrix}\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}\\\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}\\\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&-\boldsymbol{I}&\mathbf{0}\\\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&-\boldsymbol{I}\\\mathbf{0}&\mathbf{0}&\boldsymbol{I}&\mathbf{0}&\mathbf{0}&\mathbf{0}\\\mathbf{0}&\mathbf{0}&\mathbf{0}&\boldsymbol{I}&\mathbf{0}&\mathbf{0}\end{pmatrix},$$

$$\mathcal{K}_i=\begin{pmatrix}\mathbf{0}&\mathbf{0}&\mathbf{0}&\frac{1}{2}\mathcal{R}_i&\mathbf{0}&\mathbf{0}\\\mathbf{0}&\mathbf{0}&-\frac{1}{2}\mathcal{R}_i&\mathbf{0}&\mathbf{0}&\mathbf{0}\\\mathbf{0}&-\frac{1}{2}\mathcal{R}_i&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}\\\frac{1}{2}\mathcal{R}_i&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}\\\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}\\\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}\end{pmatrix}, \tag{8.2.9}$$

而 $\mathcal{M}$ 中的元素 $\boldsymbol{I}$ 指 3 阶单位矩阵, $\mathcal{K}_i$ 中的 $\mathcal{R}_i$ $(i=1,2,3)$ 是一个 3 阶反对称矩阵:

$$\mathcal{R}_1=\begin{pmatrix}0&0&0\\0&0&-1\\0&1&0\end{pmatrix},\quad\mathcal{R}_2=\begin{pmatrix}0&0&1\\0&0&0\\-1&0&0\end{pmatrix},\quad\mathcal{R}_3=\begin{pmatrix}0&-1&0\\1&0&0\\0&0&0\end{pmatrix}. \tag{8.2.10}$$

$\mathcal{M}$ 显然是反对称的. 每个 $\mathcal{K}_i$ 都是一个 18×18 的方阵, 它们也是反对称的. 另外, 旋度算子为

$$\nabla\times=\mathcal{R}_1\frac{\partial}{\partial x}+\mathcal{R}_2\frac{\partial}{\partial y}+\mathcal{R}_3\frac{\partial}{\partial z}.$$

多辛 Bridges 形式也就是多辛 Hamilton 形式, 由 Bridges 等人首次提出. 我们在第六章已将它推广到最一般的多辛 Birkhoff 形式. 多

辛 Hamilton 形式有下面的守恒律:

$$\frac{\partial \omega}{\partial t} + \nabla \times \boldsymbol{\kappa} = 0, \tag{8.2.11}$$

其中

$$\omega = \frac{1}{2}\mathrm{d}\overline{\boldsymbol{Z}} \wedge \boldsymbol{M}\mathrm{d}\overline{\boldsymbol{Z}}, \quad \mathrm{d}\overline{\boldsymbol{Z}} = (\mathrm{d}\overline{Z}_1, \cdots, \mathrm{d}\overline{Z}_{2n})^{\mathrm{T}}, \tag{8.2.12}$$

$$\boldsymbol{\kappa} = \left(\frac{1}{2}\mathrm{d}\overline{\boldsymbol{Z}} \wedge \mathcal{K}_1\mathrm{d}\overline{\boldsymbol{Z}}, \frac{1}{2}\mathrm{d}\overline{\boldsymbol{Z}} \wedge \mathcal{K}_2\mathrm{d}\overline{\boldsymbol{Z}}, \frac{1}{2}\mathrm{d}\overline{\boldsymbol{Z}} \wedge \mathcal{K}_3\mathrm{d}\overline{\boldsymbol{Z}}\right)^{\mathrm{T}} \triangleq (\kappa_x, \kappa_y, \kappa_z)^{\mathrm{T}}.$$

对任意两个 n 维向量函数 $\boldsymbol{p} = (p_1, \cdots, p_n)^{\mathrm{T}}$ 和 $\boldsymbol{q} = (q_1, \cdots, q_n)^{\mathrm{T}}$, 外积符号 $\wedge$ 表示

$$\mathrm{d}\boldsymbol{p} \wedge \mathrm{d}\boldsymbol{q} = \sum_{i=1}^{n} \mathrm{d}p_i \wedge \mathrm{d}q_i. \tag{8.2.13}$$

系统 (8.2.7) 有局部多辛守恒律

$$\frac{1}{2}\nabla \times (\mathrm{d}\boldsymbol{U} \wedge \mathrm{d}\boldsymbol{H} - \mathrm{d}\boldsymbol{V} \wedge \mathrm{d}\boldsymbol{E}) - \frac{\partial}{\partial t}(\mathrm{d}\boldsymbol{V} \wedge \mathrm{d}\boldsymbol{P} + \mathrm{d}\boldsymbol{U} \wedge \mathrm{d}\boldsymbol{Q}) = 0. \tag{8.2.14}$$

现在我们再回过头去看上一节讨论过的 Maxwell 方程的第一种形式. 同样, 它要求 ε 和 μ 不依赖于时间变量 t, 但不要求 $\varepsilon = \mu$. 它可以直接写成多辛 Hamilton 形式

$$\begin{aligned}
&\begin{pmatrix} \boldsymbol{0} & -\varepsilon\boldsymbol{\partial}_t \\ \mu\boldsymbol{\partial}_t & \boldsymbol{0} \end{pmatrix}\begin{pmatrix} \boldsymbol{H} \\ \boldsymbol{E} \end{pmatrix} + \begin{pmatrix} \mathcal{R}_1\partial_x & \boldsymbol{0} \\ \boldsymbol{0} & \mathcal{R}_1\partial_x \end{pmatrix}\begin{pmatrix} \boldsymbol{H} \\ \boldsymbol{E} \end{pmatrix} \\
&\quad + \begin{pmatrix} \mathcal{R}_2\partial_y & \boldsymbol{0} \\ \boldsymbol{0} & \mathcal{R}_2\partial_y \end{pmatrix}\begin{pmatrix} \boldsymbol{H} \\ \boldsymbol{E} \end{pmatrix} \\
&\quad + \begin{pmatrix} \mathcal{R}_3\partial_z & \boldsymbol{0} \\ \boldsymbol{0} & \mathcal{R}_3\partial_z \end{pmatrix}\begin{pmatrix} \boldsymbol{H} \\ \boldsymbol{E} \end{pmatrix} = \begin{pmatrix} \boldsymbol{J} \\ -\boldsymbol{K} \end{pmatrix},
\end{aligned} \tag{8.2.15}$$

这里协变 Hamilton 量为 $S = \langle \boldsymbol{J}, \boldsymbol{H} \rangle - \langle \boldsymbol{K}, \boldsymbol{E} \rangle$. 形式 (8.2.15) 的局部多辛守恒律为

$$\begin{aligned}
&\frac{\partial}{\partial x}(\mathcal{R}_1\mathrm{d}\boldsymbol{H} \wedge \mathrm{d}\boldsymbol{H} + \mathcal{R}_1\mathrm{d}\boldsymbol{E} \wedge \mathrm{d}\boldsymbol{E}) + \frac{\partial}{\partial y}(\mathcal{R}_2\mathrm{d}\boldsymbol{H} \wedge \mathrm{d}\boldsymbol{H} + \mathcal{R}_2\mathrm{d}\boldsymbol{E} \wedge \mathrm{d}\boldsymbol{E}) \\
&\quad + \frac{\partial}{\partial z}(\mathcal{R}_3\mathrm{d}\boldsymbol{H} \wedge \mathrm{d}\boldsymbol{H} + \mathcal{R}_3\mathrm{d}\boldsymbol{E} \wedge \mathrm{d}\boldsymbol{E}) - \frac{\partial}{\partial t}(\mathrm{d}\boldsymbol{E} \wedge \mathrm{d}\boldsymbol{H}) = 0.
\end{aligned} \tag{8.2.16}$$

§8.3 电磁场方程的多辛算法

本节我们考虑多辛 Hamilton 形式 (8.2.6) 的多辛算法. 因为系统 (8.2.6) 包括 18 个方程, 是一个比较大的方程组, 如果考虑 1+3 维的情况, 计算量太大, 所以我们这里只给出了 1+2 维多辛格式的构造, 并且假设外加磁场强度为零, 即 $\boldsymbol{K}=\mathbf{0}$. 对于 $\boldsymbol{K}\neq\mathbf{0}$ 的情况, 格式的构造几乎一样. 此时旋度算子变为

$$\nabla\times\overline{\boldsymbol{Z}}=\boldsymbol{\mathcal{R}}_1\overline{\boldsymbol{Z}}_x+\boldsymbol{\mathcal{R}}_2\overline{\boldsymbol{Z}}_y. \tag{8.3.1}$$

我们用中点格式同时沿 x 方向, y 方向和 t 方向离散方程组 (8.2.6), 得到

$$\begin{aligned}
&\boldsymbol{\mathcal{R}}_1\frac{\boldsymbol{U}^{k+\frac12}_{i+1,j+\frac12}-\boldsymbol{U}^{k+\frac12}_{i,j+\frac12}}{2h_x}+\boldsymbol{\mathcal{R}}_2\frac{\boldsymbol{U}^{k+\frac12}_{i+\frac12,j+1}-\boldsymbol{U}^{k+\frac12}_{i+\frac12,j}}{2h_y}\\
&\quad=\boldsymbol{P}^{k+\frac12}_{i+\frac12,j+\frac12}-\mu_{i+\frac12,j+\frac12}\boldsymbol{H}^{k+\frac12}_{i+\frac12,j+\frac12},\\
&-\boldsymbol{\mathcal{R}}_1\frac{\boldsymbol{V}^{k+\frac12}_{i+1,j+\frac12}-\boldsymbol{V}^{k+\frac12}_{i,j+\frac12}}{2h_x}-\boldsymbol{\mathcal{R}}_2\frac{\boldsymbol{V}^{k+\frac12}_{i+\frac12,j+1}-\boldsymbol{V}^{k+\frac12}_{i+\frac12,j}}{2h_y}\\
&\quad=\boldsymbol{Q}^{k+\frac12}_{i+\frac12,j+\frac12}-\varepsilon_{i+\frac12,j+\frac12}\boldsymbol{E}^{k+\frac12}_{i+\frac12,j+\frac12},\\
&-\frac{\boldsymbol{P}^{k+1}_{i+\frac12,j+\frac12}-\boldsymbol{P}^{k}_{i+\frac12,j+\frac12}}{\tau}-\boldsymbol{\mathcal{R}}_1\frac{\boldsymbol{E}^{k+\frac12}_{i+1j+\frac12}-\boldsymbol{E}^{k+\frac12}_{i,j+\frac12}}{2h_x}\\
&\qquad-\boldsymbol{\mathcal{R}}_2\frac{\boldsymbol{E}^{k+\frac12}_{i+\frac12,j+1}-\boldsymbol{E}^{k+\frac12}_{i+\frac12,j}}{2h_y}=\mathbf{0},\\
&-\frac{\boldsymbol{Q}^{k+1}_{i+\frac12,j+\frac12}-\boldsymbol{Q}^{k}_{i+\frac12,j+\frac12}}{\tau}\\
&\qquad+\boldsymbol{\mathcal{R}}_1\frac{\boldsymbol{H}^{k+\frac12}_{i+1,j+\frac12}-\boldsymbol{H}^{k+\frac12}_{i,j+\frac12}}{2h_x}+\boldsymbol{\mathcal{R}}_2\frac{\boldsymbol{H}^{k+\frac12}_{i+\frac12,j+1}-\boldsymbol{H}^{k+\frac12}_{i+\frac12,j}}{2h_y}\\
&\quad=\boldsymbol{J}^{k+\frac12}_{i+\frac12,j+\frac12},\quad\frac{\boldsymbol{V}^{k+1}_{i+\frac12,j+\frac12}-\boldsymbol{V}^{k}_{i+\frac12,j+\frac22}}{\tau}=\boldsymbol{H}^{k+\frac12}_{i+\frac12,j+\frac12},\\
&\frac{\boldsymbol{U}^{k+1}_{i+\frac12,j+\frac12}-\boldsymbol{U}^{k}_{i+\frac12,j+\frac12}}{\tau}=\boldsymbol{E}^{k+\frac12}_{i+\frac12,j+\frac12},
\end{aligned}\tag{8.3.2}$$

其中下标 i 和 j 分别代表 x 方向和 y 方向的格点位置, k 代表时间方向的格点位置. 上面的格式称为 **Central box 格式** (也称为 Preissman 格式), 它在相应的离散空间上有如下离散的多辛守恒律:

$$\frac{\kappa_{xi+1,j+\frac{1}{2}}^{k+\frac{1}{2}}-\kappa_{xi,j+\frac{1}{2}}^{k+\frac{1}{2}}}{2h_x}+\frac{\kappa_{yi+\frac{1}{2},j+1}^{k+\frac{1}{2}}-\kappa_{yi+\frac{1}{2},j}^{k+\frac{1}{2}}}{2h_y}-\frac{\omega_{i+\frac{1}{2},j+\frac{1}{2}}^{k+1}-\omega_{i+\frac{1}{2},j+\frac{1}{2}}^{k}}{\tau}=0, \tag{8.3.3}$$

其中

$$\begin{aligned}
\kappa_{xi,j+\frac{1}{2}}^{k+\frac{1}{2}}&=\mathcal{R}_1\mathrm{d}\boldsymbol{U}_{i,j+\frac{1}{2}}^{k+\frac{1}{2}}\wedge\mathrm{d}\boldsymbol{H}_{i,j+\frac{1}{2}}^{k+\frac{1}{2}}-\mathcal{R}_1\mathrm{d}\boldsymbol{V}_{i,j+\frac{1}{2}}^{k+\frac{1}{2}}\wedge\mathrm{d}\boldsymbol{E}_{i,j+\frac{1}{2}}^{k+\frac{1}{2}},\\
\kappa_{yi+\frac{1}{2},j}^{k+\frac{1}{2}}&=\mathcal{R}_2\mathrm{d}\boldsymbol{U}_{i+\frac{1}{2},j}^{k+\frac{1}{2}}\wedge\mathrm{d}\boldsymbol{H}_{i+\frac{1}{2},j}^{k+\frac{1}{2}}-\mathcal{R}_2\mathrm{d}\boldsymbol{V}_{i+\frac{1}{2},j}^{k+\frac{1}{2}}\wedge\mathrm{d}\boldsymbol{E}_{i+\frac{1}{2},j}^{k+\frac{1}{2}},\\
\omega_{i+\frac{1}{2},j+\frac{1}{2}}^{k}&=\mathrm{d}\boldsymbol{V}_{i+\frac{1}{2},j+\frac{1}{2}}^{k}\wedge\mathrm{d}\boldsymbol{P}_{i+\frac{1}{2},j+\frac{1}{2}}^{k}+\mathrm{d}\boldsymbol{U}_{i+\frac{1}{2},j+\frac{1}{2}}^{k}\wedge\mathrm{d}\boldsymbol{Q}_{i+\frac{1}{2},j+\frac{1}{2}}^{k}.
\end{aligned}\tag{8.3.4}$$

从 Central box 格式消去引进的变量 $\boldsymbol{V},\boldsymbol{U},\boldsymbol{P}$ 和 $\boldsymbol{Q}$, 我们得到原电磁场方程的一个多辛格式

$$\begin{aligned}
&\frac{\tau}{h_x}\mathcal{R}_1\Big(\boldsymbol{E}_{i+2,j+2}^{k+1}+2\boldsymbol{E}_{i+2,j+1}^{k+1}+\boldsymbol{E}_{i+2,j}^{k+1}-\boldsymbol{E}_{i,j+2}^{k+1}-2\boldsymbol{E}_{i,j+1}^{k+1}-\boldsymbol{E}_{i,j}^{k+1}\\
&\quad+\boldsymbol{E}_{i+2,j+2}^{k}+2\boldsymbol{E}_{i+2,j+1}^{k}+\boldsymbol{E}_{i+2,j}^{k}-\boldsymbol{E}_{i,j+2}^{k}-2\boldsymbol{E}_{i,j+1}^{k}-\boldsymbol{E}_{i,j}^{k}\Big)\\
&\quad+\frac{\tau}{h_y}\mathcal{R}_2\Big(\boldsymbol{E}_{i+2,j+2}^{k+1}+2\boldsymbol{E}_{i+1,j+2}^{k+1}+\boldsymbol{E}_{i,j+2}^{k+1}-\boldsymbol{E}_{i+2,j}^{k+1}-2\boldsymbol{E}_{i+1,j}^{k+1}-\boldsymbol{E}_{i,j}^{k+1}\\
&\quad+\boldsymbol{E}_{i+2,j+2}^{k}+2\boldsymbol{E}_{i+1,j+2}^{k}+\boldsymbol{E}_{i,j+2}^{k}-\boldsymbol{E}_{i+2,j}^{k}-2\boldsymbol{E}_{i+1,j}^{k}-\boldsymbol{E}_{i,j}^{k}\Big)\\
&=\mu\Big(\boldsymbol{H}_{i+2,j+2}^{k+1}+2\boldsymbol{H}_{i+2,j+1}^{k+1}+2\boldsymbol{H}_{i+1,j+2}^{k+1}+4\boldsymbol{H}_{i+1,j+1}^{k+1}\\
&\quad+\boldsymbol{H}_{i,j+2}^{k+1}+2\boldsymbol{H}_{i,j+1}^{k+1}+\boldsymbol{H}_{i+2,j}^{k+1}+2\boldsymbol{H}_{i+1,j}^{k+1}+\boldsymbol{H}_{i,j}^{k+1}\Big)\\
&\quad-\mu\Big(\boldsymbol{H}_{i+2,j+2}^{k}+2\boldsymbol{H}_{i+2,j+1}^{k}+2\boldsymbol{H}_{i+1,j+2}^{k}+4\boldsymbol{H}_{i+1,j+1}^{k}\\
&\quad+\boldsymbol{H}_{i,j+2}^{k}+2\boldsymbol{H}_{i,j+1}^{k}+\boldsymbol{H}_{i+2,j}^{k}+2\boldsymbol{H}_{i+1,j}^{k}+\boldsymbol{H}_{i,j}^{k}\Big),
\end{aligned}\tag{8.3.5}$$

$$\begin{aligned}
&\frac{\tau}{h_x}\mathcal{R}_1\Big(\boldsymbol{H}_{i+2,j+2}^{k+1}+2\boldsymbol{H}_{i+2,j+1}^{k+1}+\boldsymbol{H}_{i+2,j}^{k+1}-\boldsymbol{H}_{i,j+2}^{k+1}-2\boldsymbol{H}_{i,j+1}^{k+1}-\boldsymbol{H}_{i,j}^{k+1}\\
&\quad+\boldsymbol{H}_{i+2,j+2}^{k}+2\boldsymbol{H}_{i+2,j+1}^{k}+\boldsymbol{H}_{i+2,j}^{k}-\boldsymbol{H}_{i,j+2}^{k}-2\boldsymbol{H}_{i,j+1}^{k}-\boldsymbol{H}_{i,j}^{k}\Big)\\
&\quad+\frac{\tau}{h_y}\mathcal{R}_2\Big(\boldsymbol{H}_{i+2,j+2}^{k+1}+2\boldsymbol{H}_{i+1,j+2}^{k+1}+\boldsymbol{H}_{i,j+2}^{k+1}-\boldsymbol{H}_{i+2,j}^{k+1}-2\boldsymbol{H}_{i+1,j}^{k+1}-\boldsymbol{H}_{i,j}^{k+1}\\
&\quad+\boldsymbol{H}_{i+2,j+2}^{k}+2\boldsymbol{H}_{i+1,j+2}^{k}+\boldsymbol{H}_{i,j+2}^{k}-\boldsymbol{H}_{i+2,j}^{k}-2\boldsymbol{H}_{i+1,j}^{k}-\boldsymbol{H}_{i,j}^{k}\Big)
\end{aligned}$$

$$
\begin{aligned}
&= -\varepsilon\Big(\boldsymbol{E}_{i+2,j+2}^{k+1} + 2\boldsymbol{E}_{i+2,j+1}^{k+1} + 2\boldsymbol{E}_{i+1,j+2}^{k+1} + 4\boldsymbol{E}_{i+1,j+1}^{k+1} \\
&\quad + \boldsymbol{E}_{i,j+2}^{k+1} + 2\boldsymbol{E}_{i,j+1}^{k+1} + \boldsymbol{E}_{i+2,j}^{k+1} + 2\boldsymbol{E}_{i+1,j}^{k+1} + \boldsymbol{E}_{i,j}^{k+1}\Big) \\
&\quad + \varepsilon\Big(\boldsymbol{E}_{i+2,j+2}^{k} + 2\boldsymbol{E}_{i+2,j+1}^{k} + 2\boldsymbol{E}_{i+1,j+2}^{k} + 4\boldsymbol{E}_{i+1,j+1}^{k} \\
&\quad + \boldsymbol{E}_{i,j+2}^{k} + 2\boldsymbol{E}_{i,j+1}^{k} + \boldsymbol{E}_{i+2,j}^{k} + 2\boldsymbol{E}_{i+1,j}^{k} + \boldsymbol{E}_{i,j}^{k}\Big).
\end{aligned}
$$

我们给一个简单的算例来检验格式 (8.3.5). 分别取 $\varepsilon = \mu = 1$ 和 $\boldsymbol{J} = \boldsymbol{0}$, 并取初始条件为

$$
\begin{aligned}
E_1(x,y,0) &= -\frac{\sqrt{2}}{2}\cos(2\pi x)\sin(2\pi y), \\
E_2(x,y,0) &= \frac{\sqrt{2}}{2}\sin(2\pi x)\cos(2\pi y), \\
H_3(x,y,0) &= \sqrt{\frac{\varepsilon}{\mu}}\cos(2\pi x)\cos(2\pi y), \\
H_1(x,y,0) &= H_2(x,y,0) = E_3(x,y,0) = 0,
\end{aligned}
\tag{8.3.6}
$$

在这个条件下, 方程的 Poynting 能量是守恒的, 即有

$$
\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t} = 0, \quad \mathcal{H} = \iint_{(x,y)} (\varepsilon|\boldsymbol{E}|^2 + \mu|\boldsymbol{H}|^2)\mathrm{d}x\mathrm{d}y, \tag{8.3.7}
$$

同时方程有一个二次局部能量守恒律

$$
\frac{\partial}{\partial t}(\varepsilon|\boldsymbol{E}|^2 + \mu|\boldsymbol{H}|^2) + 2\nabla\cdot(\boldsymbol{E}\times\boldsymbol{H}) = 0. \tag{8.3.8}
$$

我们的计算区域取为 $(x,y)\in[0,1]\times[0,1]$. x 方向和 y 方向取同样的步长: $h_x = h_y = 0.04$, 时间步长取为 $\tau = 0.02$. 图 8.3.1(a), (b) 显示的是 H_3 在时刻 $t = 100\tau$ 的波形, 其中 (a) 为精确解, (b) 为数值解.

我们还计算了 0 到 1000τ 时间段内整体能量守恒 (8.3.7) 和局部能量守恒 (8.3.8) 的误差, 分别在图 8.3.2 和图 8.3.3 中显示. 计算结果说明, 格式 (8.3.5) 同时保持局部能量和整体能量守恒.

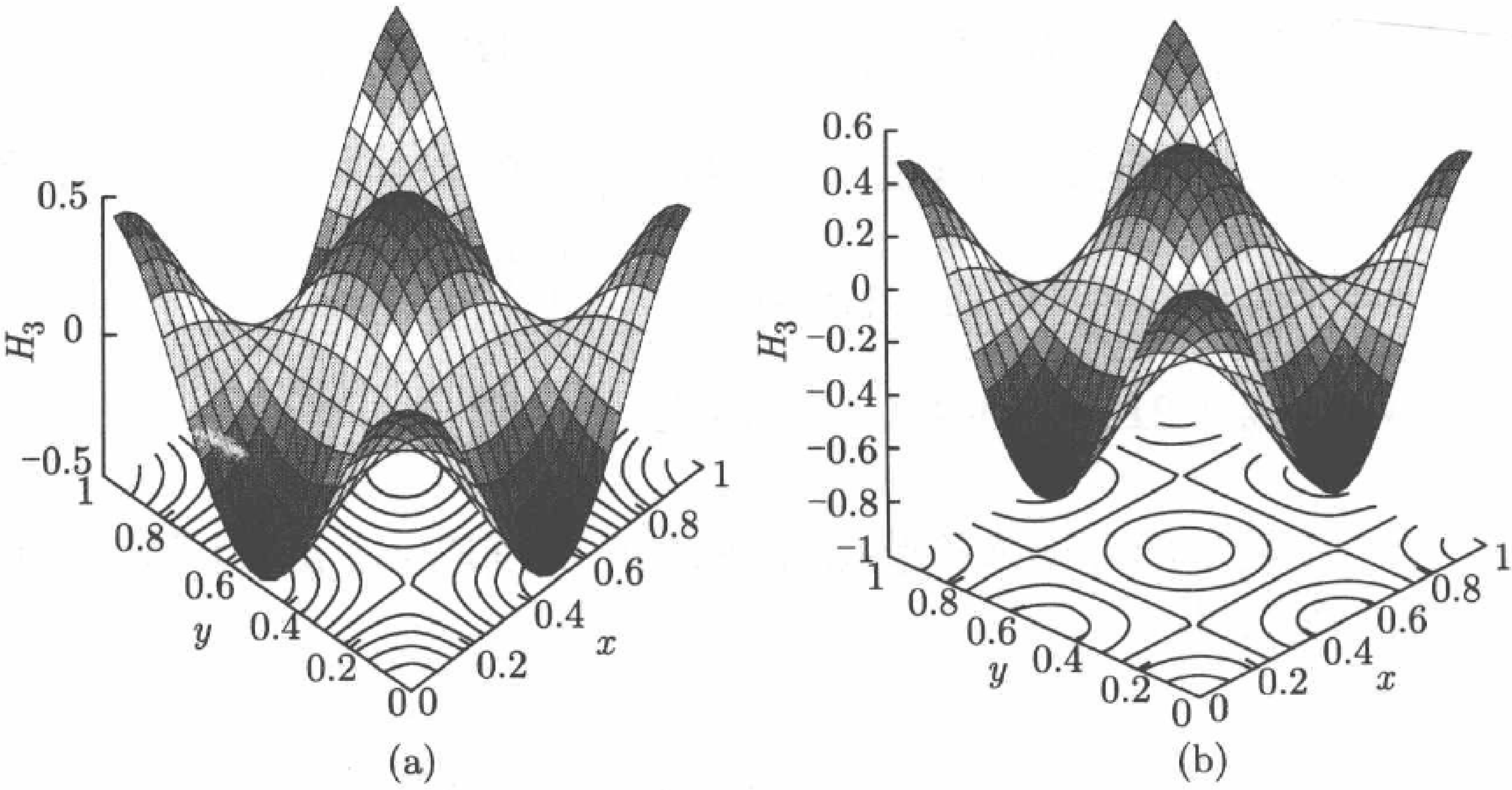


图 8.3.1　磁场波形

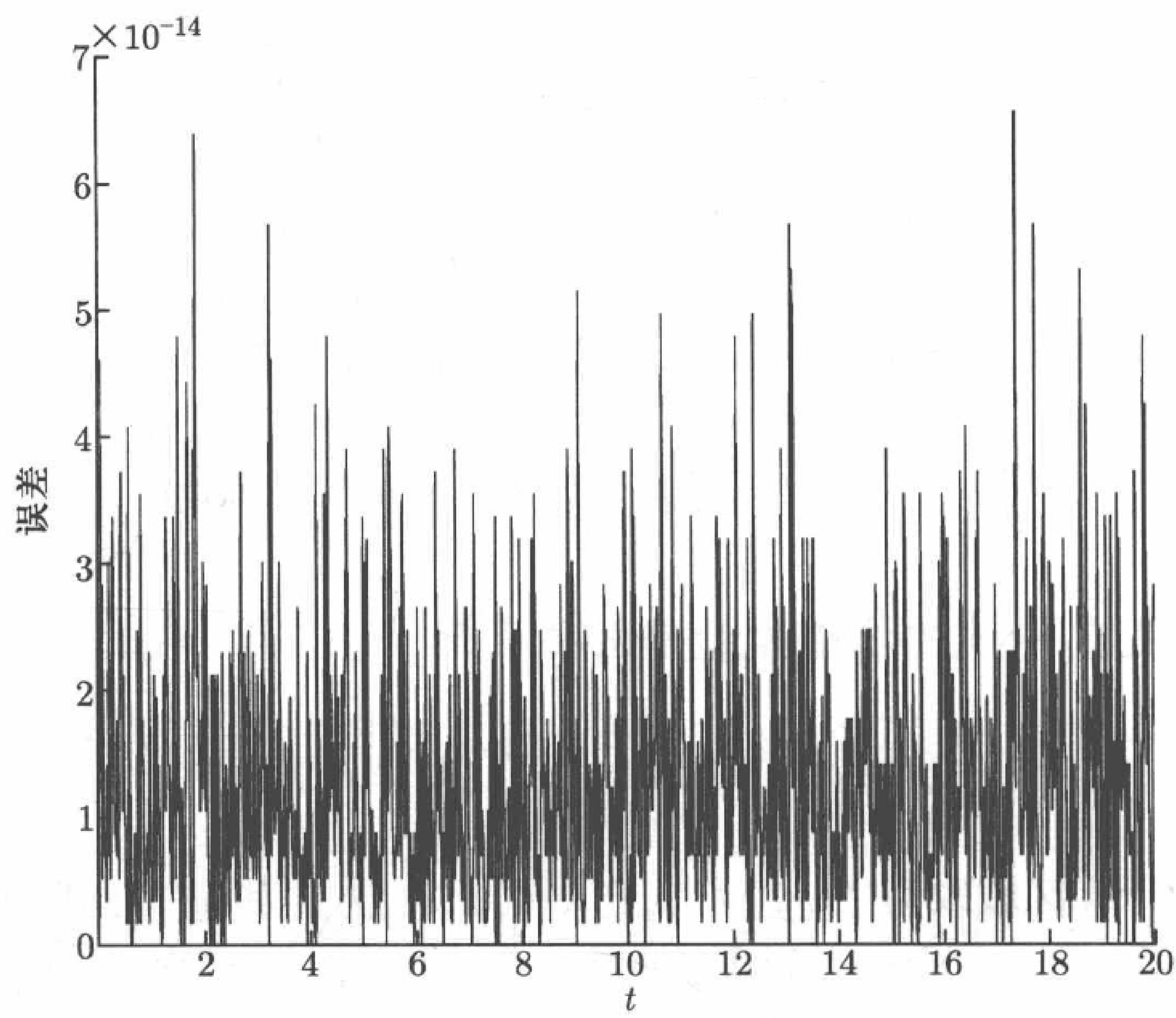


图 8.3.2　整体能量守恒误差

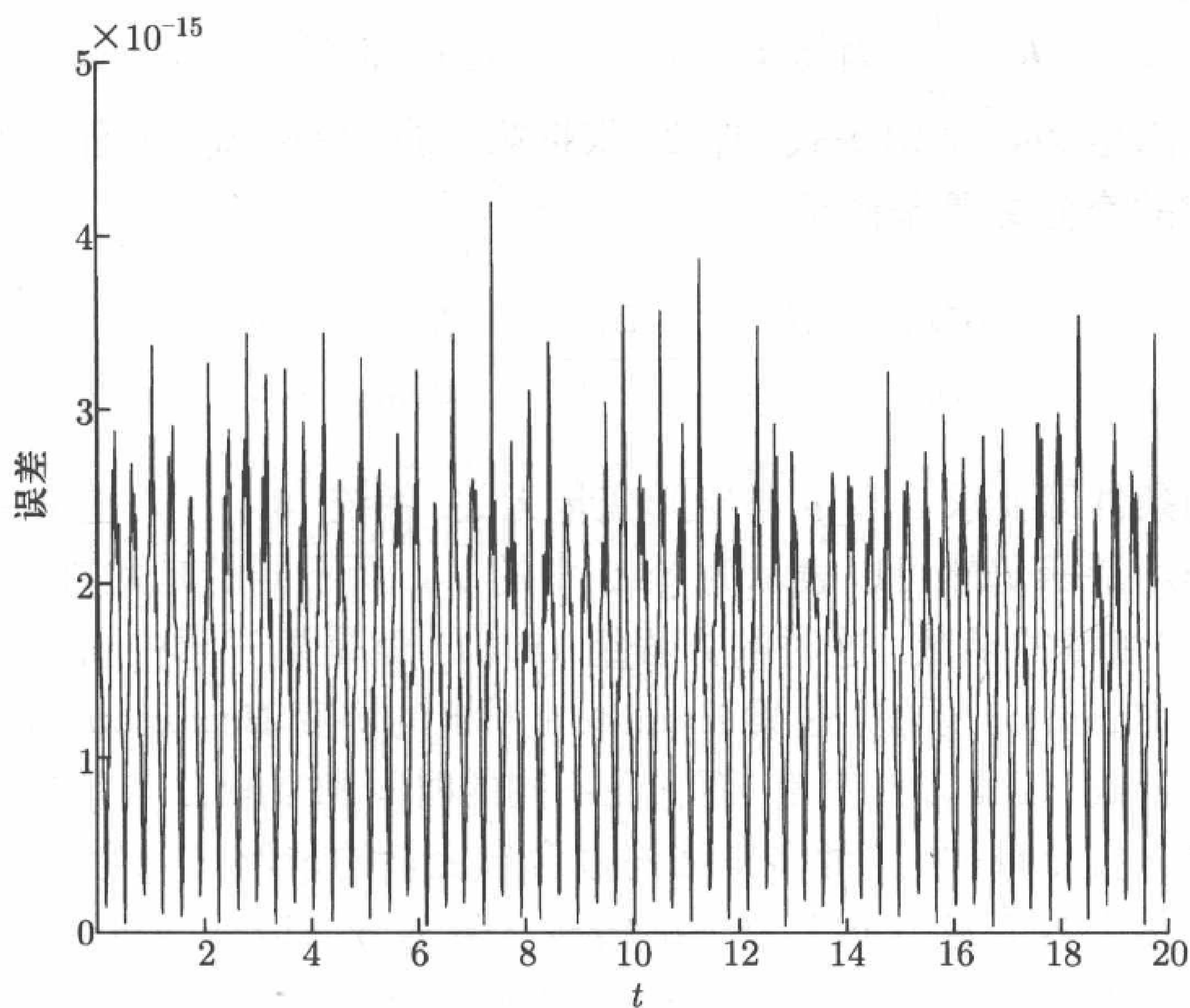


图 8.3.3 局部能量守恒误差

§8.4 更一般的 Maxwell 方程的多辛 Hamilton 表示

假设外加的电流密度和磁流密度是随时间变化的向量函数, 即

$$\boldsymbol{J} = \boldsymbol{J}(x, y, z, t), \quad \boldsymbol{K} = \boldsymbol{K}(x, y, z, t),$$

则 Maxwell 方程可以描述为

$$\boldsymbol{G}^t \boldsymbol{Z}_t + \boldsymbol{G}^x \boldsymbol{Z}_x + \boldsymbol{G}^y \boldsymbol{Z}_y + \boldsymbol{G}^z \boldsymbol{Z}_z - \boldsymbol{F} = \boldsymbol{0}. \tag{8.4.1}$$

我们取

$$\begin{aligned} \boldsymbol{G}^t &= \begin{pmatrix} \boldsymbol{0} & -\boldsymbol{\partial}_t \\ \boldsymbol{\partial}_t & \boldsymbol{0} \end{pmatrix}, \quad \boldsymbol{G}^x = \begin{pmatrix} a\boldsymbol{\mathcal{R}}_1 & \boldsymbol{0} \\ \boldsymbol{0} & b\boldsymbol{\mathcal{R}}_1 \end{pmatrix}, \\ \boldsymbol{G}^y &= \begin{pmatrix} a\boldsymbol{\mathcal{R}}_2 & \boldsymbol{0} \\ \boldsymbol{0} & b\boldsymbol{\mathcal{R}}_2 \end{pmatrix}, \quad \boldsymbol{G}^z = \begin{pmatrix} a\boldsymbol{\mathcal{R}}_3 & \boldsymbol{0} \\ \boldsymbol{0} & b\boldsymbol{\mathcal{R}}_3 \end{pmatrix}, \end{aligned} \tag{8.4.2}$$

其中 $a=\dfrac{1}{\varepsilon}$, $b=\dfrac{1}{\mu}$. 因为 $\boldsymbol{G}^{\alpha}\,(\alpha=t,x,y,z)$ 都不显含 $\boldsymbol{Z}$, 所以以上四个矩阵满足 Jacobi 恒等式. 那么, 根据第三章中的结论, 方程组 (8.4.1) 是对称的充分必要条件是

$$\frac{\partial G^{x}_{\beta\nu}}{\partial x}+\frac{\partial G^{y}_{\beta\nu}}{\partial y}+\frac{\partial G^{z}_{\beta\nu}}{\partial z}=\frac{\partial F_{\beta}}{\partial Z_{\nu}}-\frac{\partial F_{\nu}}{\partial Z_{\beta}}. \tag{8.4.3}$$

上面的条件即要求 ε 和 μ 与空间变量无关. 系统 (8.4.1) 对应的变分问题的作用泛函仍然可以表示为 §8.2 中的 L_2, 所不同的是此时的 $\boldsymbol{J}$, $\boldsymbol{K}$ 是时–空变量的函数. 这时的方程组 (8.4.1) 就是第三章中指出的特殊的多辛 Hamilton 系统. 这类 Hamilton 系统的协变 Hamilton 函数不显含未知函数, 仅显含自变量, 而它在不同的时间和空间方向上的辛结构却是显含自变量的, 这种多辛结构显含自变量的情形又分为两种情况: 一种是每个自变量方向上的辛结构显含除本方向以外的自变量; 另一种是各个方向的辛结构因为显含自变量而产生的耗散项正好互相抵消了. 所以, 这里若 ε 和 μ 不依赖于空间变量, 那么 Maxwell 方程 (8.4.1) 就属于第一种特殊的 Hamilton 系统. 假如 ε 和 μ 依赖于空间变量, 那么只要它们可以表示为

$$\varepsilon=\frac{f(x,y,z)}{a(t)},\quad \mu=\frac{f(x,y,z)}{b(t)}, \tag{8.4.4}$$

则 Maxwell 方程就属于第二种特殊的 Hamilton 系统, 它仍然可用表达式 (8.4.1) 表示, 其中

$$\begin{aligned}
\boldsymbol{G}^{t}&=\begin{pmatrix}\mathbf{0} & -f(x,y,z)\boldsymbol{\partial}_t\\ f(x,y,z)\boldsymbol{\partial}_t & \mathbf{0}\end{pmatrix},\\
\boldsymbol{G}^{\alpha}&=\begin{pmatrix}a(t)\boldsymbol{\mathcal{R}}_{\alpha} & \mathbf{0}\\ \mathbf{0} & b(t)\boldsymbol{\mathcal{R}}_{\alpha}\end{pmatrix},\quad \alpha=x,y,z.
\end{aligned} \tag{8.4.5}$$

从这个多辛形式出发我们就可以研究方程组的多辛几何以及多辛算法了. 当 ε,μ 以及外加电场和磁场都与 $\boldsymbol{H}$ 和 $\boldsymbol{E}$ 有关时, 也可以类似讨论对应的电磁场方程的 Hamilton 多辛结构和 Birkhoff 多辛结构.

参 考 文 献

[Yee66] Yee K S. Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Trans Amtennas Propagation*, 1966, 14: 302–307.

[Taflove75] Taflove A, Brodwin M E. Numerical solution of steady-state electromagnetic scattering problems using the time-dependent Maxwell's equations. *IEEE Trans Microw Theory Tech*, 1975, 23: 623–630.

[Taflove05] Taflove A, Hagness S. *Computational Electromagnetics: The Finite-Difference Time-Domain Method.* 3rd edition, Boston: Artech House, 2005.

[SQ03] Su Hong-Ling, Qin Meng-Zhao. Multisymplectic geometry method for Maxwell's equations and multisymplectic scheme. *arXiv: math-ph*/0302058, 2003, 24.

[SQ07] Su H L, Qin M Z, Scherer R. A multisymplectci geometry and a multisymplectic scheme for Maxwell's equations. *Int J Pure Appl Math*, 2007, 34: 1–17.

[KH10] Kong L, Hong J, Zhang J. Splitting multisymplectic integrators for Maxwell's equations. *J Comput Phys*, 2010, 229: 4259–4278.

[SunT11] Sun Y, Tse P S P. Symplectic and multisymplectic numerical methods for Maxwell's equations. *J Comput Phys*, 2011 230: 2076–2094.

[CWS13] Cai W, Wang Y, Song Y. Numerical dispersion analysis of a multi-symplectic scheme for the three dimensional Maxwell's equations. *J Comput Phys*, 2013, 234: 330–352.

[GZH13] 龚丽红, 张泽, 郑宇, 吴非. 基于离散变分原理的Birkhoff 辛算法. 科技创新导报, 2013, 19: 210–211.

[KWM12] Kong Xin-lei, Wu Hui-bin, Mei Feng-xiang. Structure-preserving algorithms for autonomous Birkhoffian systems. *Journal of Beijing Institute of Technology*, 2012, 21(1): 1–7.

[ZHW99] Zhu Haiping, Wu Jike. Generalized cononical transformations

and symplectic algorithm of the autonomous Birkhoffian systems. *Prog Nat Sci*, 1999, 9(11): 820–828.

[KWM13] Kong Xinlei, Wu Huibin, Mei Fengxiang. Discrete optimal control for Birkhoffian systems. *Nonlin Dyn*, 2013, DOI 10.1007/s11071-013-0999-0.

[LSG13] 刘畅, 宋瑞, 刘世兴, 郭永新. 非齐次 Hamilton 系统的 Birkhoff 表示. 结构和系统动力学与控制专刊, 2013, 43(4): 541–548.

[XPZL12] 解加芳, 庞硕, 邹杰涛, 李国富. 非完整系统 Boltzmann-Hamel 方程的 Birkhoff 化及其广义辛算法. 物理学报, 2012, 61(23): 1–5.

[ZHW02] Zhang Xing-wu, Wu Ji-ke, Zhu Hai-ping, Huang Ke-fu. Generalized cononical transformation and symplectic algorithms of the autonomous Birkhoffian systems. *App Math Mech*, 2002, 23(9): 1029–1034.

[LLG11] Liu Shi-Xing, Liu Chang, Guo Yong-Xin. Geometric formulations and variational integrators of discrete autonomous Birkhoff systems. *Chin Phys B*, 2011, 20(3): 034501.

[KWM12] Kong X, Wu H, Mei F. Structure-preserving algorithms for Birkhoffian systems, *J Geom Phys*, 2012, 60: 1157–1166.